

Coffman, J. (2003, June). *Lessons in evaluating communications campaigns: Five case studies*. Cambridge, MA: Harvard Family Research Project. Builds on the findings of the first and second papers. It examines specifically how campaigns with different purposes (individual behavior change and policy change) have been evaluated, and how evaluators have tackled some of the associated evaluation challenges that the first three papers raised as important to address. It features five brief case studies in which the main unit of analysis is not the campaign, but the campaign's *evaluation*. The case studies provide a brief snapshot of the real experiences of campaign evaluations. The paper also features cross-case lessons that highlight important findings and themes.

LESSONS IN EVALUATING COMMUNICATIONS CAMPAIGNS

FIVE CASE STUDIES

Prepared for the Communications Consortium Media Center

By Julia Coffman
Harvard Family Research Project
June 2003

I. INTRODUCTION AND BACKGROUND.....	1
II. TWO MAIN CAMPAIGN TYPES.....	2
III. CAMPAIGN THEORIES OF CHANGE.....	3
A. Individual Behavior Change Campaigns	4
B. Policy Change Campaigns	5
IV. EVALUATION CASE STUDIES.....	7
A. Safe Gun Storage Campaign	8
B. Stop It Now! Vermont	13
C. Voluntary Ozone Action Program	18
D. Community Trials Project	23
E. The Devolution Initiative	30
V. CROSS-CASE STUDY LESSONS.....	36
REFERENCES.....	39
APPENDIX A.....	42

The author thanks George Perlov, The Advertising Council; Joan Tabachnick, Stop it Now! Vermont; Gary Henry, Georgia State University; Andrew Treno, Prevention Research Center; and the W.K. Kellogg Foundation for their willingness to share their experiences for the case studies in this paper. The author also thanks Gary Henry, Georgia State University, and Heather Weiss, Harvard Family Research Project, for their comments on earlier drafts of this paper.

I. INTRODUCTION AND BACKGROUND

This is the fourth in a series of short-term research projects and papers that are part of a larger movement coordinated by the Communications Consortium Media Center to build capacity for improved evaluation in the nonprofit communications arena (see Appendix A for a description of all four projects). Each of these projects, including this one, was designed to take us closer to the goal of responding to the acute need in this field for better information about *how* to evaluate communications efforts. The need for practical advice is great, including examples of how others have already responded to the distinct evaluation challenges that communications efforts present.

Previous research and the opinions of experts about the needs and opportunities in the field of communications evaluation have revealed some key findings (e.g., Bohan-Baker, 2001b; Coffman, 2002b; Dorfman, Ervice, & Woodruff, 2002; Dungan-Seaver, 1999; Henry, 2002; Sutton & Thompson, 2002):

Evaluation practice needs to keep pace with campaign advancements. As the practice of “strategic communications” has grown in recent years and proliferated among nonprofits that recognize its potential value, interest in determining whether and how these practices add value and how they can be improved for better effect also has grown. But the actual practice of, and knowledge base about, nonprofit communications evaluation simply has not increased at a concomitant or satisfactory rate – both in terms of the number of people who evaluate their communications efforts and the quality of the evaluation work that is being done. This leaves many wondering whether increased investments in communications have been worth it.

We need to know more about evaluating campaigns to change policy. Most of us are familiar with communications campaigns for the purpose of individual behavior change (e.g., stop smoking, don’t drink and drive). These types of campaigns of course are still common, but another type of nonprofit campaign has taken popular hold in recent years – campaigns for the purpose of achieving policy change. Yet while their popularity is growing, we know little about, and practice less, how to evaluate efforts designed to cause policy change.

A certain standard of rigor should be expected. While different evaluation needs and questions require different standards of methodological rigor, communications evaluation in general needs to focus on more than just what is easy to count. We have to focus on more than just the outputs or measures of our own communications performance (e.g., counting the number of op-eds written or press conferences held), and push ourselves to focus more on the outcomes of those efforts for the audiences we are trying to reach.

We need to keep evaluation design options open and focus more on learning. While we need to raise the bar on our methodological rigor in general, at the same time we have to remember that not all evaluation needs to conform to the “gold standard” of research and evaluation design – experimental designs that use random assignment to control and experimental groups or conditions. Sometimes the need is not to determine whether the communications effort definitively *caused* later effects, but to learn how to do the work better. This means opening the range of evaluation possibilities up to more than just the traditional research paradigm in which the evaluation remains separate from the campaign and does not contribute to

learning or continuous improvement. This does not mean evaluators cannot be objective or rigorous in their evaluation approach. It merely means evaluation designs need to be more in tune with the evaluation questions being raised, and more open (regardless of the design used) to feeding back information that is going to be useful and used.

This paper is intended to move us closer to developing the tools we need to address these issues and take advantage of the possibilities that lay before us. It presents five case studies of already-completed campaign evaluations. They represent a diverse array of both campaign types and evaluation approaches, and offer specific examples of how evaluation can respond to the challenges summarized above. Each case study includes a set of lessons about what can be learned and applied from the evaluation, and the paper finishes with a set of cross-case-study lessons gleaned from these evaluations and others.

Note that this paper uses the term “communications campaign” broadly. Communications efforts are not always called campaigns. The case studies in this paper demonstrate that they can also be labeled programs, projects, or initiatives. In addition, campaigns do not have to be stand-alone entities, nor do they have to be highly formal efforts. As the case studies here show, in fact, very few communications campaigns are stand-alone efforts. They can also be an organized set of communication activities that are embedded within, or complementary to, a larger set of work designed to achieve a common end (Dorfman et al., 2002). While this last fact presents evaluation challenges, it also presents unique opportunities for learning about what it takes to make social change happen.

II. TWO MAIN CAMPAIGN TYPES

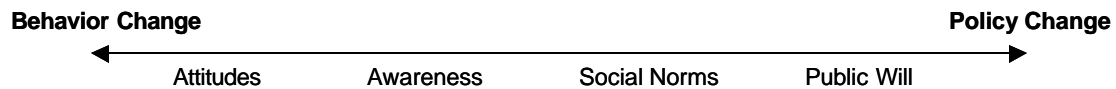
Although communication campaigns come in many varieties and cannot be placed into clean mutually-exclusive categories, the field is increasingly making a distinction between at least two main types of campaigns based on their primary purpose, or what the campaigns are ultimately trying to achieve: *individual behavior change* versus *policy change* campaigns (e.g., Dorfman et al., 2002; Dungan-Seaver, 1999; Henry & Rivera, 1998). The case studies in this paper feature both campaign types.¹

Individual behavior change campaigns try to decrease in individuals the behaviors that lead to social problems or promote behaviors that lead to improved individual or social well-being. Policy change campaigns attempt to mobilize public and decision maker support for policy support or change (Coffman, 2002b; Dorfman et al., 2002; Dungan-Seaver, 1999; Weiss & Tschirhart, 1994).

In some ways, as is illustrated in the case studies, campaigns for both end purposes look similar. For example, the types of activities used are often the same (e.g., dissemination of informational materials, public service announcements, etc.). So are some of the variables campaigns try to affect on their way to achieving either purpose. For example, Figure 1 shows other variables that fall along the continuum toward each end – attitudes, awareness, social norms, and public will.

¹ Case studies feature three individual behavior change campaigns and two policy change campaigns.

Figure 1. Continuum of Campaign Types According to Purpose



While affecting public will, or the extent to which the public cares about and supports a specific policy position, is more often associated with policy change campaigns, behavior change campaigns may also have an element of public will in their design, as public will may be related to social pressures or norms about how an individual should or should not behave (e.g., anti-smoking campaigns).² As Ethel Klein (2002) said:

The trick is figuring out where the locus of responsibility for the behavior change should be. Is it purely on the individual performing the behavior or is it also on creating public will and social responsibility to help make that change happen? You have to look at the issue's epidemiology, and say, "What are the root causes and what sustains the behavior?"

Similarly, while behavior change campaigns are almost always associated with attempts to affect awareness and attitudes, the same may be true for policy change campaigns. Communications efforts geared toward policy change often include an element of public will building for that change. If that will does not already exist, then those efforts may make an attempt to build it by influencing awareness or attitudes about that policy issue and position.

The difference in a campaign's primary purpose (individual behavior versus policy change) is a critical factor in decisions about how a campaign's evaluation will be designed. On the one hand, determining a campaign's purpose is important because it affects how a campaign will be judged in the end on the degree to which it has been successful. On the other hand, and more importantly, purpose is important to evaluation because it is the primary factor that drives a campaign's theory of change, or what an evaluation is ultimately designed around. The next section discusses theories of change that underlie campaigns for each purpose and illustrates the types of relationships between the variables that commonly appear between campaign activities and their end results.

III. CAMPAIGN THEORIES OF CHANGE

Theories of change force us to think through and put down on paper what we are doing (our activities) in connection to what we are trying to achieve (our outcomes), and to lay out the pathways and variables through which we expect change (behavior or policy) to occur. They are a representation of what we think needs to be in place to make change happen. Typically a theory of change is based on a combination of objective evidence, theory, experience, and subjective opinion and personal ideology (Frumkin, 2002). The theory is not fixed, but may change and improve as the theory is tested and more evidence gathered. A discussion about the theories of change that underlie communications campaigns is important to any discussion about evaluation because they can be a critical driving force in decision making about an evaluation's design.

² The terms policy change campaigns and public will campaigns are often used interchangeably.

Each of the five case studies featured here includes the campaign's theory of change as an introduction to the evaluation approach used.³ While each theory of change is unique, campaigns for similar purposes (individual behavior change or policy change) tend to share some commonalities. As such, what follows is an attempt to lay out general theories of change for individual behavior and policy change campaigns, which then serve as an introduction to the thinking behind the case study theories of change that appear later. They are based on theoretical and empirical research, as well as an examination of many campaign models. They are not the *only* theories of change that campaigns are based on, as the case studies in this paper reveal. Rather, they represent a composite and simplified sketch of the types of variables and relationships that commonly link communications activities with their intended results, i.e., individual behavior change and policy change.

A. Individual Behavior Change Campaigns

Individual behavior change campaigns try to change in individuals the behaviors that lead to social problems or promote behaviors that lead to improved individual or social well-being. Well-known campaigns in this category target behaviors such as smoking, drug use, designated driving, seat belt usage, domestic violence, or fire and crime prevention. Many come from the public health arena, but this type of campaign has branched out into other areas such as education and early childhood (see Dorfman et al., 2002, for a description of numerous behavior change campaigns).

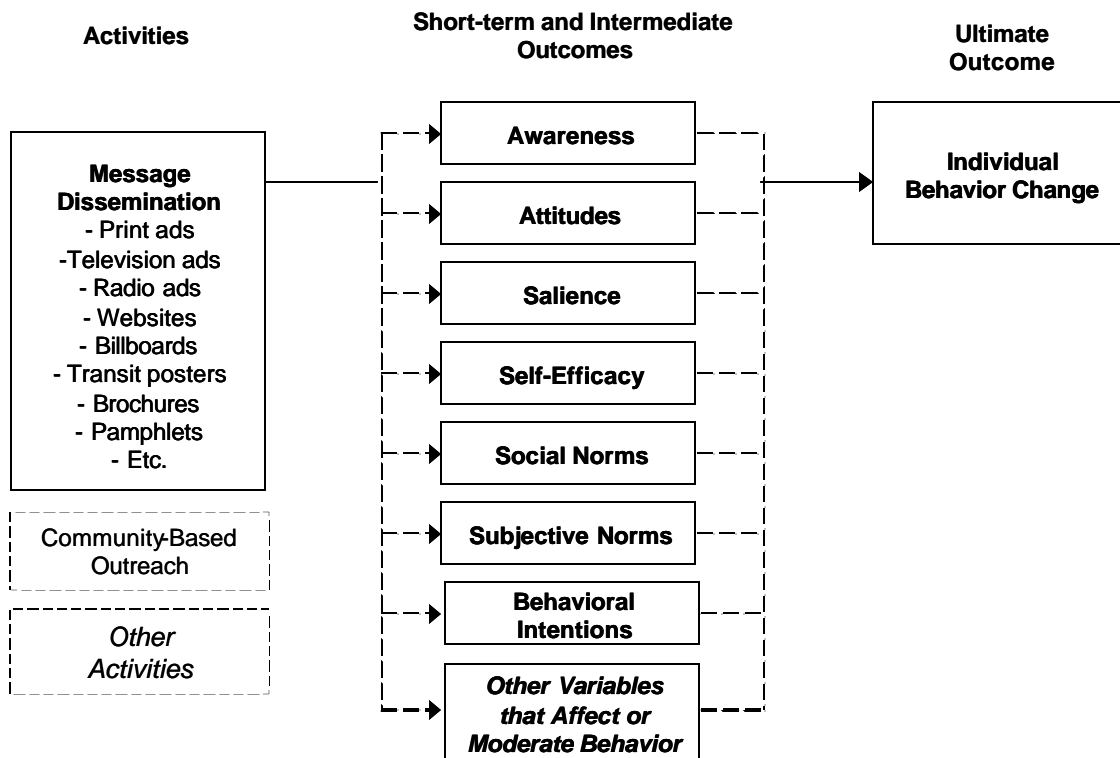
Fortunately, decades of behavior change research have taught us much about what it takes to affect human behavior. In the last decade in particular, much progress has been made on incorporating social science theory into both campaign design and evaluation (see, for example, the Voluntary Ozone Action Program case study in this paper). Behavior change theorists and researchers now agree on a number of factors that have been proven to influence behavior (Fishbein, Triandis, Kanfer, Becker, Middlestadt, & Eichler, 2001), and campaign designers have benefited from this knowledge. For example, research and experience tells us that information to increase a person's knowledge or awareness about the need to change or adopt a particular behavior typically does not change behavior on its own. How many people are still smoking even though they know the substantial cancer risks? As a result, fewer campaigns are only disseminating information to increase people's awareness. More are designed to influence other aspects of how we think about and act on issues, such as whether we have the self-efficacy (perception in our capability to perform the behavior) necessary to change the behavior, or what our perceptions are about what our friends and family are doing (social norms) or want us to do (subjective norms).

Figure 2 offers a general theory of change for behavior change campaigns. The first column represents broadly the various types of message dissemination a campaign might use. This is not a comprehensive list. That column also shows boxes with dashed lines to represent the fact that such efforts are usually implemented alongside other activities also designed to contribute to the ultimate outcome of behavior change. While those activities would presumably have relationships with at least some of the outcomes

³ The author for purposes of illustration in this paper developed these figures. Campaign documents, materials, conceptual models, and logic models informed them.

represented here (and others of their own), this model addresses only the relationships between communications-related activities and their outcomes.

Figure 2. General Theory of Change for Individual Behavior Change Campaigns



The short-term and intermediate outcomes in this model include a list of common variables that communications activities typically try to affect, and that we know from research can have a relationship with behavior change (see also Coffman, 2002a). This also is not a comprehensive list. Note that a communications effort typically does not try to affect *all* of the variables represented in the model. Rather, the specific variables that apply to a campaign will vary. For example with some campaigns, particularly either mature campaigns or campaigns in which the issues or behaviors being addressed are already well-known, increasing awareness will not be a necessary outcome. That campaign may focus more on trying to get people to care more about performing a behavior or on their attitudes toward performing that behavior.

The model also does not represent the specific order and pathways between the activities and outcomes. For example, some campaigns view awareness as a shorter-term outcome or precursor to attitude change. Others would insert saliency between those two variables, arguing that a person has to care about the behavior or feel it is relevant in order for attitudes to change. The case studies illustrate the different ways of representing these middle variables and relationships.

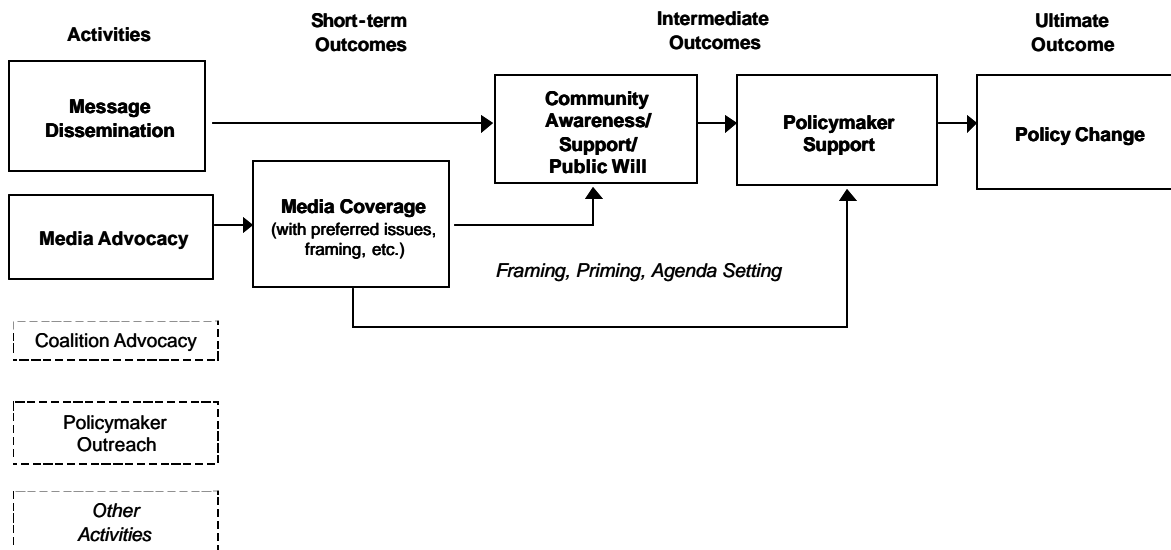
B. Policy Change Campaigns

Policy change campaigns attempt to mobilize public and decision maker support for policy change. These types of campaigns are less understood than individual behavior

change campaigns, but as they increase rapidly in number, so does our need to understand how they work (see also Salmon, Post, & Christensen, 2003 for more about these kinds of campaigns).

Figure 3 presents a general theory of change behind policy change campaigns. Like the model offered for behavior change campaigns, this is not the only theory of change on which such campaigns are based. Rather, it represents a composite sketch of common types of variables and relationships between those variables that begin with campaign (and other) activities and end with policy change (and implementation). While this model focuses on communications activities that are designed to affect policy change, it acknowledges the fact that rarely do communications activities alone achieve policy change. Typically they act as a complement to other policy advocacy activities, such as coalition or community organizing, or one-on-one policymaker outreach. These additional activities are represented as boxes with dashed lines in the activity column of Figure 3.

Figure 3. General Theory of Change for Policy Change Campaigns



The activity of *media advocacy* assumes a prominent position in this model, as it is common to many communications efforts working toward policy change. Media advocacy is the strategic use of news media to advance a social or public policy initiative (Holder & Treno, 1997; Wallack & Dorfman, 2001). Activities central to media advocacy include 1) elaborating policy options that are to be supported, 2) identifying the policymakers with power over decision making on those options, 3) identifying the audiences that can be organized to put pressure on policymakers, and 4) targeting the audiences with messages in support of the policy options (Wallack & Dorfman, 2001).

Media advocacy is designed to influence both the quantity and quality of media coverage about a particular issue. It presumes that the way an article or news story is framed – the way information is organized, the pictures used, sources referenced, etc. – has an effect on the way people perceive the issue being discussed (Gamson, 2000; Ryan, 1991). As such, it focuses on issues as *social* rather than personal problems; the idea is to get many people to care about and act on an issue, not just those with the problems.

Media advocacy often also presumes that people will pay more attention to and care about an issue if it is placed in a local context. Therefore it uses activities that “capture” the media’s attention, like local events, rallies, or press conferences that the media can attend and cover. Media advocacy can also include developing relationships with reporters or editorial boards that may influence coverage, helping reporters put national news or issues in a local context, or writing editorials and submitting op-eds.

The effect of increased and/or better media coverage as a result of media advocacy is theorized to be two-fold. It is expected to: 1) influence community awareness, support, and public will, which then in turn can impact policymaker support (whether or not the policymaker personally cared about or supported the issue in the first place), and/or 2) impact policymaker support directly.

Media advocacy can garner public and policymaker support through various theoretical mechanisms that may include agenda setting, framing, or priming. With *agenda setting*, the media does not tell people what to think, it tells them what they should think *about* and which issues are important. Thus the more often an issue appears in the news, the more important it appears (McCombs & Shaw, 1973).⁴ With *framing*, the media packages information in ways that affect people’s perceptions of particular issues (Bales, 2002; Bohan-Baker, 2001a). *Priming* holds that the media’s attention to some issues and not others alters the standards by which people evaluate issues, people, or objects. Priming assumes that people do not have a lot of knowledge about a lot of things, and therefore make decisions based on what comes to mind first, which is often what they saw or read in the media (Iyengar & Kinder, 1987).⁵

Again, the ultimate outcome here is policy change, which theoretically should have a direct relationship with policymaker support for a campaign’s stance on a particular issue.

IV. EVALUATION CASE STUDIES

This section presents a set of five evaluation case studies that are designed to serve as illustrations for how to evaluate organized communications efforts. Thus far the discussion about communications evaluation has not been tied together or grounded in comprehensive examples. These case studies offer those examples. They are campaigns that have already faced and dealt with difficult evaluation choices and challenges. Each chose a different evaluation approach, and each adds to our learning about conducting evaluation in this field. Along with their diversity in terms of evaluation approach, the five campaigns were chosen for their diversity in terms of their purpose and target audience.

Each case study also features these five elements:

- *Theory of change* – a visual representation and description of how each campaign was designed to achieve its ultimate purpose and the short and intermediate outcomes it intended to achieve along the way

⁴ Although if the issue appears too often in the media, the public may get bored with it.

⁵ See also Salmon, et al., 2003 for other theoretical mechanisms that may come into play.

- *Focus/Methods* – the focus of the evaluation or what outcomes were assessed in the theory of change, and the methods used to measure outcomes
- *Design/Analysis* – the overall design and analysis used to elicit findings
- *Key Findings* – a brief synopsis of key findings from the evaluation
- *Evaluation Lessons* – what can be learned from the evaluation experience and applied, along with select advantages and disadvantages of the approach used.

A. Safe Gun Storage Campaign ⁶

Purpose	Behavior Change	↔	Policy Change
Target Audience	General Public	Targeted Public	Policy makers
Evaluation	Experimental	Quasi-Experimental	Other/ Non-Experimental

Eight children and teens die each day from gunfire in America. Many of these deaths are accidental and could have been prevented if parents had kept their guns stored in a safe place and out of children's reach (Children's Defense Fund, 2002). Yet of the 35-40% of U.S. households that own at least one gun, up to one-third of all handguns are stored unlocked and loaded in the owner's home or vehicle, and more than half are stored unlocked (Advertising Council, 2002).

Responding to these alarming statistics, the U.S. Department of Justice, National Crime Prevention Council, and the Advertising (Ad) Council developed the Safe Gun Storage public education campaign to encourage gun owners to store their guns safely (unloaded, with both ammunition and gun locked away). The campaign ran from Summer 2000 through Fall 2002, and was distributed nationwide to a target audience of gun-owning parents with children in the home. It included television, radio, print, interactive, and out-of-home (e.g., billboard, transit posters) public service messages.

Ads were based on pre-campaign research about the attitudes and behaviors of gun-owning parents. That research found that gun owners felt danger was unpredictable and random, and that their homes (unlike schools and neighborhoods), were the last safe haven for their children. They also felt that guns were a way of maintaining that safe haven. Many parents therefore connected guns in the house with safety, rather than as a potential harm to their children. As a result, some had not thought through safe gun storage issues. Compounding this were feelings of defensiveness in general about gun ownership given the political and social climate related to gun control issues, and strong beliefs in the constitutional right to bear firearms.

Based on this research, the campaign's strategy was not to question parents' reasons for owning guns, but to get them to understand that gun ownership can pose hazards to

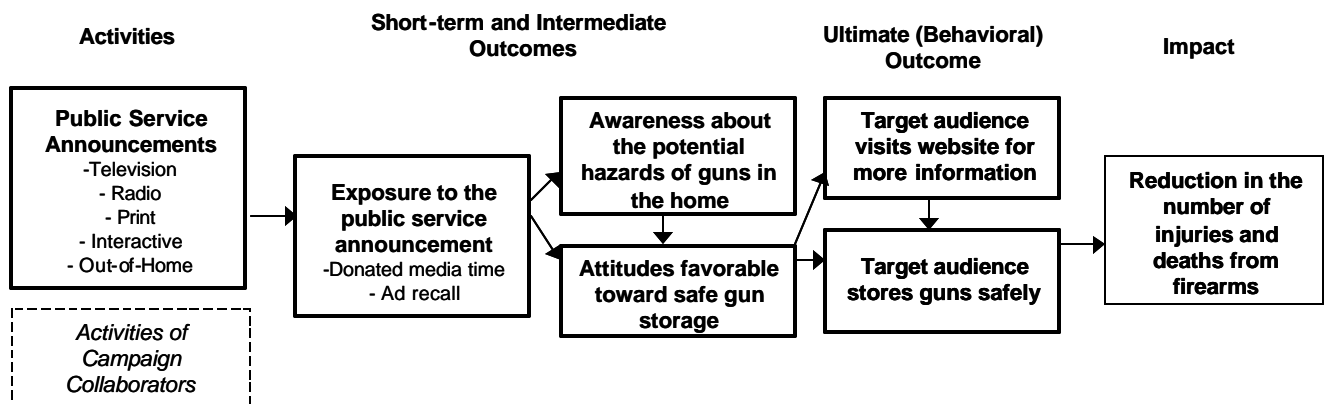
⁶ This case study was informed by telephone interviews and informal communication with George Perlov, Senior Vice President and Director of Planning and Research at the Advertising Council. It was also informed by a document review of campaign and evaluation materials, including Advertising Council (in press).

their children, and to inform them the importance of how to store guns safely. Ads used actual crayon drawings and voiceovers from children trying to cope with accidental shootings in their homes in order to bring to light the child's perspective and to challenge the entrenched framing that guns promoted safety (rather than presented danger) in the home.

Theory of Change

Figure 4 shows the Safe Gun Storage campaign's basic theory of change. Campaign activities primarily involved the release of public service announcements in various media markets. Also included on the model, as connected to but not formally a part of the campaign, are the activities of the campaign sponsors and collaborators (e.g., the National Crime Prevention Council), who at the same time that the campaign ran were using other activities to work toward the same ultimate outcomes and impact.

Figure 4. Safe Gun Campaign Theory of Change



The first short-term outcome in the model is exposure to the public service announcements, the idea being that the audience had to first be exposed to the ads in order for them to have any effects. Exposure is then linked to two additional outcomes – awareness and attitudes. Because pre-campaign research showed that some parents were not connecting gun ownership with potential hazards to their children, the campaign tried to raise awareness of those hazards. In addition, the campaign was trying to make parents' attitudes about safe gun storage more favorable, and less defensive.

More favorable attitudes were then expected to link to behavior change, or actual safe gun storage. If more parents started storing their guns safely, the desired impact was anticipated to be a decrease in the number of children injured or killed in accidental firearm incidents. The campaign also encouraged individuals to seek more information about safe gun storage on a website linked to the campaign.

Focus/Methods

The Ad Council managed the evaluation, but contracted with outside advertising research firms to conduct it. These firms provided their services at a reduced cost, as is typical with Ad Council campaign evaluations. The firms included Michael Scavone, Inc., Ipsos-ASI, and Elrick and Lavidge.

The evaluation focused on the parts of the theory of change that are bolded in Figure 4. These included 1) the quality of the ads produced, and audience 2) exposure to the ads, 3) awareness of the potential hazards of unsafe storage, 4) attitudes about safe gun storage, and 5) behavior change in terms of safe gun storage.

The *quality of campaign materials* produced was assessed by a campaign review committee of leading advertising experts and by examining audience reactions to them – whether the target audience understood the main messages, perceived their clarity, and felt they were relevant and persuasive. Michael Scavone, Inc., an independent research firm, conducted this “communications check” using interviews with target audience members after they viewed various ads. Interview formats were similar to small focus groups, and included half-hour triads (three respondents), twenty-minute one-on-one interviews, and half-hour couples interviews.

Quality of television ads was also assessed with a quantitative “copy test” to gather more data on target audience responses. The purpose was to determine ad recall, relevance, and persuasiveness. The advertising research company Ipsos-ASI conducted the copy test using their in-home, in-program methodology, which meant contacting randomly selected target audience members by telephone, and telling them they had been chosen to view a new television program. They were then sent a video of a television show, with one of the ads embedded. After participants saw the video, they were contacted again and interviewed about their responses to the ad. Measures included *measured attention* – the percentage of respondents who remembered the ad when read an unbranded description of it, and *brand linkage* – the ratio of related recall (linking the brand with the ad) and measured attention. The copy test also included a “control” condition in which responses of the target audience were compared with responses to the general population.

Exposure was assessed by calculating the dollar value of donated media time for each of the two years that ads were available. Dollar values were calculated separately for donated media time from broadcast media, cable television, radio, newspaper, magazine, out-of-home placements, Internet, and other media.⁷ The Ad Council contracted with various media tracking services to assess donated media time and assign dollar values.⁸

The remaining outcomes – *awareness*, *attitudes*, and *behavior*, were assessed using three target audience surveys with repeated measures – one implemented before the campaign, one during, and one after two years had passed. Elrick and Lavidge, a marketing research firm, conducted the surveys. The pre-campaign survey took place in June 2000 with a random sample of 500 adults over 18 years of age living in households with children, and with someone in the household owning a gun. The second survey, implemented one year after the campaign’s launch (June 2001), was conducted with a similar random sample. The third survey was conducted after the second year of the campaign (July 2002),⁹ again with a similar random sample.¹⁰ The evaluation also tracked hits to the website linked to the campaign as a behavioral indicator.

⁷ Ad Council campaigns are distributed using donated media time, as opposed to paid media time. Media outlets have to choose to air or print the PSA; the campaign has little control itself over that decision.

⁸ See Perlov (2002) for a description of tracking services the Ad Council uses.

⁹ The campaign is still in the field, however, no new materials are being developed.

¹⁰ Note that the surveys selected different samples of individuals for each implementation.

Design/Analysis

The Safe Gun Storage campaign evaluation examined differences at the three measurement time points (random surveys of the target population) in terms of awareness, attitude, and behavior. This design did not allow for any definitive conclusions about causal links between the campaign and shifts in awareness, attitudes, or behavior measured from the first survey through the third, but it did allow for comparisons of where the target audience was on these three variables of interest over time.¹¹

Key Findings

About campaign quality:

- Both the communications check and copy test methods revealed that the target audience found the ads were persuasive and were not turned off by them. Measures of ad recall during the copy test were high.

About campaign exposure:

- Over two years, according to estimates using the method described above, the ads received over \$65 million worth of donated media across all media formats.

About awareness and attitudes:

- Awareness about safe gun storage ads was high throughout the campaign.¹² Surveys revealed that over half of the respondents indicated they had seen a PSA about safe gun storage; this percentage changed little over the two years.
- The percentage of gun owners who defined safe storage as “locking” their guns in a cabinet or rack increased from 38% in the first survey to 58% in the third.
- The percentage of gun owners who felt it was important to lock up guns in households with children remained unchanged over time.

About behavior:

- Small differences in reported gun storage behavior were noted across the three survey waves. The percentage of gun owners who reported moving their guns over the previous six months increased from 6% in the second survey to 12% at the end of the campaign. The Advertising Council concluded that a large percentage of gun owners were already storing their guns safely, but that the campaign helped to move that percentage higher.
- Other behaviors showed no changes over time, such as the percentage of gun owners who kept loaded weapons in the home (12%), and those who kept their guns unloaded, but keep the ammunition in the same room (two-thirds).

Evaluation Lessons

When many people think about campaigns designed to achieve social change, they think about the Ad Council. Many non-Ad-Council campaigns, in fact, share a lot of the same communication characteristics as Ad Council campaigns – development of ads for different types of media outlets, and intended effects on audience awareness, attitudes, and behavior. As such, this case study presents a good overview of the evaluation challenges and opportunities that many campaigns face.

¹¹ While some might classify it as non-experimental, Cook, Campbell, & Peracchio (1990) define this design as quasi-experimental – an untreated control group design with independent pretest and samples.

¹² Note that other gun safety PSAs were also running at the same time, including ads from Americans for Gun Safety and the PAX “Ask” campaign.

Front-end research is a good guide for evaluation. A discussion of “front-end” research, or research that takes place to inform the development of the campaign, is not the focus of this paper. However, it is important to note the investment that the Ad Council puts into front-end research for all of its campaigns (see Perlov, 2002, for more on the Ad Council’s approach). The lesson here is that this research, and what is learned about the variables that seem to be driving the particular behavior in question, should inform the evaluation. If front-end research finds, for example, that awareness of an issue is already high, but behaviors are not changing, then awareness may not need to be a focus of the campaign or the evaluation.

Choices about evaluation designs require realistic tradeoffs. This campaign presents a classic evaluation design challenge prevalent among many campaigns – that for various reasons it is difficult to set up an evaluation that will definitively say whether the campaign caused its intended effects. For this campaign, and with most Ad Council campaigns, the primary back-end evaluation approach was to compare results from surveys taken at different time points with different groups of randomly selected target audience members. This approach is common in communications evaluation.

The Ad Council acknowledges that it is difficult to gauge the success of a public service campaign in general, and that there are limitations to what can be concluded from this type of evaluation design. For the Ad Council, controlled designs are difficult because their campaigns rely on donated media time; it is not possible to control when the intervention occurs, how often it occurs, or who is exposed to it. Therefore setting up comparison or control groups or conditions is difficult to impossible. Another factor is available resources in terms of dollars and time. These factors limit methodological choices. The Ad Council’s approach is typically less costly and returns results faster than designs that use control or comparison groups.¹³

Interpretation should be qualified with non-controlled evaluation designs. The main challenge to the Safe Gun Storage campaign’s evaluation approach is that it puts constraints on the ability to draw conclusions about campaign effects because it makes it difficult to deal with various challenges to the interpretive validity of those conclusions (Cook & Campbell, 1979). However, acknowledging those limitations, if the sample is randomly selected each time (and of an adequate sample size), it is possible to compare results from one survey to the next. When drawing comparisons with this design, however, this question needs to be considered – Would the samples or results have changed over time in ways that might not be attributable to the campaign (Cook, Campbell, & Perrachio, 1990)? In this case, the Ad Council identified and reported these factors along with results, including historical events that took place during, but were not related to, the campaign (i.e., other campaigns implemented during the same time frame; the focus on gun safety during the 2000 Presidential campaign; high-profile lawsuits against gun manufacturers about gun safety; and a steady wave of tragic shootings involving children).

¹³ This evaluation, front-end research included, cost approximately \$125,000.

B. Stop It Now! Vermont ¹⁴

Purpose	Behavior Change	↔	Policy Change
Target Audience	General Public	Targeted Public	Policy makers
Evaluation	Experimental	Quasi-Experimental	Other/Non-Experimental

As many as half a million children in this country are sexually abused each year (Villarosa, 2002). STOP IT NOW! is a national nonprofit organization founded on two core beliefs. First, that child sexual abuse can be prevented through the power of public health, and second, that children should not bear the sole burden for preventing and stopping child sexual abuse. While teaching children how to detect the difference between “good touch” and “bad touch” is important, making children the accountable

party in a differential power relationship where children are often dependent on their abusers, is a limited and incomplete preventative step. Instead, abusers should be held directly accountable for their actions. Contrasting sharply with the predominant worldview about how typical abusers behave, STOP IT NOW! calls on all abusers and those at risk to abuse to stop their behavior and seek help.

In addition to holding abusers accountable, those around the abuser are also held accountable for their actions (or lack thereof). Since 90% of the victims of child sexual abuse know their abusers, STOP IT NOW! calls on all adults – family, friends, neighbors, teachers, professionals, and abusers themselves – to prevent abuse before it happens and to report it when it does. This approach is predicated on the experience of campaigns such as those to stop drunk driving. For example, in their initial phases, Mothers Against Drunk Driving (MADD) campaigns focused on individual responsibility with the message “Don’t drink and drive.” When that approach did not do enough to reduce drunk driving, the campaign moved on to the designated driver concept. It emphasized social responsibility and messages focused on the fact that “friends don’t let friends drink and drive” (Klein, 2002). STOP IT NOW! is sending that same message of *social responsibility* for the prevention of child sexual abuse.

These beliefs are put into practice using four main categories of work. The organization develops awareness in those at risk to abuse and encourages them to seek help; offers abusers ways to seek help through a confidential helpline; works with families, peers, and friends on how to confront abusers; and works to build a social climate that does not tolerate child sexual abuse. This addresses the fact that while people do know to report abuse if they are certain it is occurring, many are unclear what to do if there is some suspicion but also some uncertainty. STOP IT NOW! provides the opportunity for people to ask in confidence about their suspicions, without reporting it or getting the police involved.

In 1995 STOP IT NOW! in partnership with the Safer Society Foundation in Vermont, created the first program design for this approach, called STOP IT NOW! VERMONT. ¹⁵

¹⁴ This case study was informed by a telephone interview and informal communication with Joan Tabachnick, STOP IT NOW! VERMONT’s Director of Public Education. It was also informed by document review of campaign and evaluation materials; Tabachnick & Dawson, (2000) was a primary source.

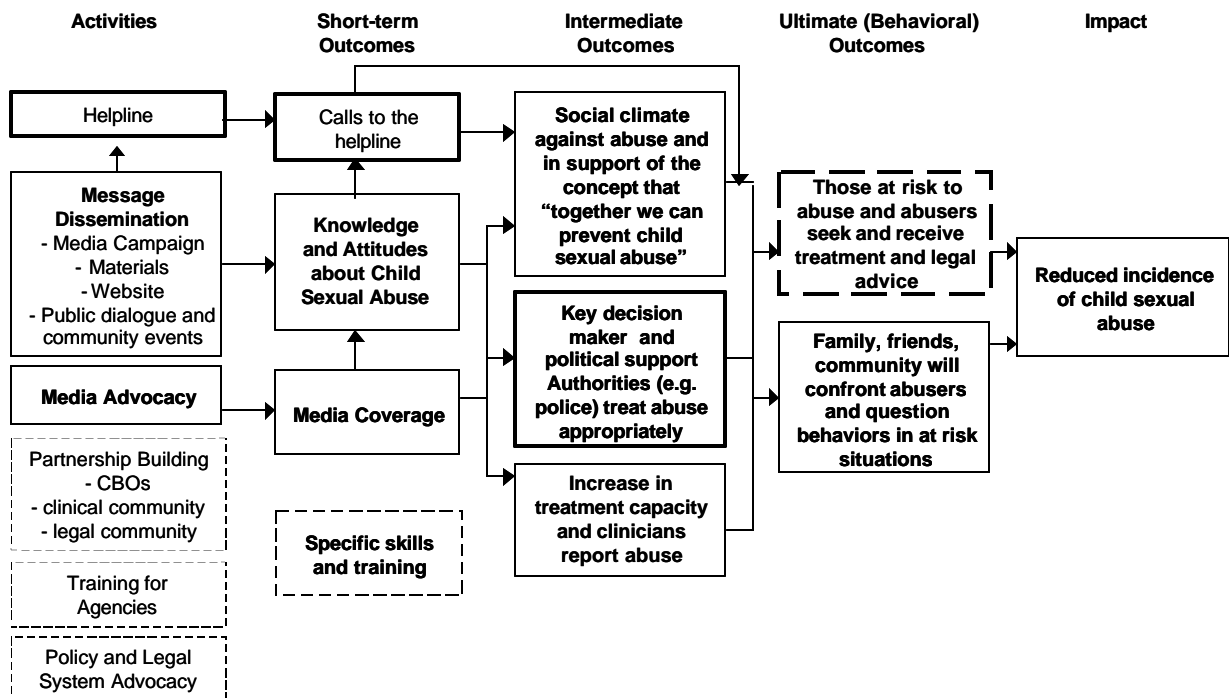
¹⁵ STOP IT NOW! organizations also exist in Haydenville, MA (national office), Philadelphia, Minnesota, and in the United Kingdom and Ireland. Note that an evaluation sponsored by the Centers for Disease Control and conducted by ORC Macro (an opinion research company), was also conducted with the Philadelphia site.

The program began as a public health social marketing campaign (it is now much more), using print ads, billboards, posters, public service announcements, a website, informational materials about child sexual abuse, and media advocacy to change the way Vermonters think about and act on this issue. The campaign is coupled with the first helpline in the country that links callers (adult or adolescent abusers, those at risk to abuse, or their friends and family) with specific Vermont resources. The campaign goals are to 1) increase public awareness of child sexual abuse, 2) challenge abusers to stop their abuse and seek treatment, 3) change attitudes about what can be done to prevent abuse, and 4) encourage behaviors by adults to intervene in abuse situations or act before the abuse occurs.

Theory of Change

Figure 5 presents the Vermont program's theory of change. The first column of activities shows the confidential helpline and the prominence of message dissemination through multiple mechanisms. This includes the media campaign, but also the dissemination of informational pamphlets and booklets, and media advocacy.

Figure 5. STOP IT NOW! Theory of Change



Also shown in Figure 5 are other activities that STOP IT NOW! VERMONT implements, which include building partnerships with other community-based organizations and the clinical and legal communities in order to build treatment capacity about how to deal effectively with this issue. It also includes a structural component of policy and legal system advocacy. This case study focuses primarily on message dissemination activities.

The first expected short-term outcome in the model is calls to the helpline. Another is change in public knowledge and attitudes about child sexual abuse. For example, in order to get the public to accept and react to their social responsibility on this issue, Stop

It Now! Vermont believes it is necessary to educate them about what exactly sexual abuse is, how prevalent it is in their communities, the early signs of detection, and what to do if they see a situation of concern, or if child sexual abuse is detected. Abusers also are educated that confidential help is available. Another short-term outcome is a change in the way the media talks about sexual abuse and abusers (e.g., not only to report on sexual abuse after the fact, but to focus on prevention, and to portray people who abuse as people who need treatment and can learn to control their behaviors).

Intermediate outcomes include a climate in which social responsibility for preventive child sexual abuse is commonly known and accepted. Currently, the norm is that if someone wants help, there appears to be nowhere safe to turn. Doctors and therapists are mandated to turn child sexual abusers in with no guarantee of help. Defense attorneys will advise individuals to keep quiet for fear of arrest, imprisonment, or community notification. Therefore the general public is not the only audience that needs to be educated on prevention and treatment options. Also included here as targets for increased awareness and capacity are key decision makers and policymakers, and the treatment and legal communities.¹⁶

Ultimate behavioral outcomes include those at risk to abuse and abusers seeking and receiving help, and those around abusers confronting them and helping them to stop their abuse. The expected impact from those ultimate outcomes is a lower incidence of child sexual abuse.

Focus/Methods

STOP IT NOW! VERMONT and the independent marketing research firm Market Street Research in Northampton, Massachusetts, collaborated to evaluate message dissemination efforts. Market Street Research provided their services in-kind and worked closely with Vermont staff on the evaluation's design and implementation. STOP IT NOW! VERMONT staff also tracked indicators of their own work, and conducted data collection within the clinical and legal community.

The evaluation focused on the bolded components of the theory of change in Figure 5. It assessed the short-term outcomes of whether calls to the helpline came in, and public knowledge, attitudes, and beliefs. The evaluation also assessed an intermediate outcome – key decision maker awareness, opinions of, and support for STOP IT NOW! VERMONT and their approach. Finally, the evaluation tested the key assumption that abusers would voluntarily seek treatment and legal advice. The latter outcome is represented with a dashed line in Figure 5 because the evaluation did not track the behavior of abusers connected to the work of STOP IT NOW! (calls to the helpline are confidential). Rather, it attempted to determine whether *any* abusers in Vermont had voluntarily sought help for their sexual behavior problems.

Helpline staff collected data about *calls to the helpline* as calls came in. The data were later compiled to determine caller characteristics, reasons for calling, and how callers heard about the helpline. This information tested the assumption that people would actually call for help.

¹⁶ Note that partnership building in the activity column would also link directly to this outcome.

Knowledge, attitudes, and beliefs were tracked with a random digit dial telephone survey of 200 Vermonters, conducted at three time points – 1995, 1997, and 1999. Because the campaign started in 1995, the first survey served as a “baseline.”

Market Street Research also conducted interviews to assess *key decision maker awareness and support* for STOP IT NOW! VERMONT, views of their successes, and the issues they felt still needed to be addressed.

Finally, STOP IT NOW! VERMONT staff collected information about *abusers seeking treatment or legal help* through a survey of clinicians and individual telephone interviews in each Vermont county prosecutor’s office.

Design/Analysis

Because it looked at outcomes in the theory-of-change separately, rather than as linked, the evaluation design did not allow for causal inferences from the campaign to its outcomes. Given the resources available and the campaign’s universal reach, an alternative design was not feasible. However, as was described in the previous Safe Gun Storage case study, this design does allow for the tracking and comparing of some key measures and indicators over time.

Key Findings

About calls to the helpline:

- People will use the helpline, including abusers and those at risk for abusing. From 1995 to 1997 STOP IT NOW! VERMONT received 657 calls to the helpline. Of those calls, 15% were from abusers and 50% were from people who knew the abuser and/or victim.
- While male callers only average about 10% on most hotlines, 32% of the callers to the Vermont helpline were male.
- The campaign is helping to get the word out about the availability of help. A little over one-half of all callers heard about the helpline from either the media or the website. Another 25% heard about it through referrals from other professionals.

About Vermonters knowledge, attitudes, and beliefs:

- Vermonters who could explain child sexual abuse almost doubled from 1995 to 1999 (45% to 85%). Overall awareness about child sexual abuse in Vermont is high and remained high over time.
- More work needs to be done in order to create the public knowledge needed to identify the warning signs for perpetration and to provide specific skills and information about what to do, especially in situations where the person is unsure if any abuse has occurred. The percentage of Vermonters who could name at least one warning sign in an adult or juvenile was low in general, but increased slightly from 1995 to 1999 (27.5% to 38%). Only a little over half of Vermonters knew where to refer someone with a sexual behavior problem compared to 77% who know where to refer someone with a drinking problem.

About key decision maker support:

- Key decision makers and stakeholders were positive about STOP IT NOW! VERMONT’s work, credibility, and positioning in the state. Interviews revealed a desire for even more involvement in and commitment to this work in Vermont.

About abusers seeking treatment or legal help:

- Abusers will voluntarily seek treatment and legal help. The survey with clinicians and telephone interviews with county prosecutor's offices identified 118 people who voluntarily sought help for sexual behavior problems (20 adults and 98 adolescents).

Evaluation Lessons

This case study featured an evaluation that was small, but resourceful. While other case studies feature much larger monetary investments to determine causal inferences, as these lessons demonstrate, this evaluation had different, but arguably no less important, payoffs.

The evaluation focused on learning. While it was operating under a very small budget,¹⁷ staff had a real commitment to *learning*, and therefore placed an emphasis on collecting data that would help them understand how to do their work better. Given that their messages were new to Vermont and that they could not invest in an evaluation model that would definitively prove causal effects in terms of behavior change, they invested in some key pieces that would help them know if some of their assumptions were correct and if they were headed in the right direction, as well as the areas in which they needed to place more or less emphasis.

The relationship between campaign staff and evaluators was productive and useful. The relationship between the external evaluator and STOP IT NOW! staff was exemplary. For example, the evaluator began working with Vermont staff in the initial stages of the campaign, and spent a lot of time working with staff on the right survey questions to ask. They later wrote their reports in ways that helped STOP IT NOW! staff learn about what areas seemed to be working and what areas needed improvement. The focus here was on making sure the evaluation information gathered was useful. As a result, STOP IT NOW! and Market Street Research have developed a lasting relationship. Their approach is an example of how to respond to the fact that too few campaign staff and evaluators form relationships that are mutually beneficial or focus on learning.

This approach shows how campaigns can collect their own evaluation data. This evaluation is also an example of how campaign staff can take responsibility on their own for identifying and tracking the information they need. This idea has not caught on as much in the communications arena as it has in the human services fields. In this case, Vermont STOP IT NOW! staff knew that it would be important to track from the very beginning information about the calls they received on their helpline. Today they can use this information to demonstrate the fact that their assumption that abusers will call is correct, and to determine what treatment and capacity needs are still unmet. Staff also took it upon themselves to conduct interviews with the treatment and legal community. While sometimes it is necessary to get an external evaluator to do data collection for purposes of objectivity, there are often many opportunities for staff to collect data that can both demonstrate effects and lead to continuous improvement.

¹⁷ Less than \$15,000, keeping in mind the polling services donated in-kind.

C. Voluntary Ozone Action Program ¹⁸

Purpose	Behavior Change	↔	Policy Change
Target Audience	General Public	Targeted Public	Policy makers
Evaluation	Experimental	Quasi-Experimental	Other/Non-Experimental

After amendments were made to the Clean Air Act of 1990, the U.S. Environmental Protection Agency (EPA) disseminated air quality standards for six airborne pollutants known to be harmful to public health, including ground-level ozone.¹⁹ The EPA also established ways of measuring compliance with those standards, and found 31 areas across the country did not meet them. One of those “non-attainment” areas was the 13-county metropolitan region around Atlanta, which specifically did not meet ground-level ozone

standards (Henry & Gordon, 2003).

As a result, in 1997, the Environmental Protection Division of Georgia’s Department of Natural Resources developed the Voluntary Ozone Action Program (VOAP) to reduce ground-level ozone in the Atlanta region. VOAP had two main components: a public information campaign and a voluntary emission reduction program.

The goals of the public information campaign were to 1) raise awareness about and the importance of ground-level ozone, 2) inform the public about the health consequences of ground-level ozone, and 3) reduce behaviors that cause harmful emissions (related to driving mostly). The campaign primarily targeted the behavior of reduced driving, which was estimated to account for 50% of Atlanta’s ground-level ozone production.

Because ozone is undetectable to human senses, it was necessary to increase community member awareness of when it was particularly critical to reduce driving. As a result, the campaign used “ozone alerts” to inform the public that ozone concentrations would be high the next day. Alerts were disseminated through electronic highway traffic signs, the local newspaper, and radio and television weather and traffic reports. The campaign also included public service announcements, news stories about ground-level ozone and pollution, and editorials about air quality.

In addition to the public information campaign, VOAP attempted to create a social context that would help lead to the desired behavioral outcomes – namely by getting employers to take ozone-reducing actions when ozone concentrations were highest (May to October and during the day-time hours). This included encouraging alternate transportation use and telecommuting among their employees, and staggering work schedules to reduce rush hour driving. For example, the Governor issued an executive order in 1997 for all state agencies, departments, and institutions of higher education to reduce the rate of single occupancy vehicles by 20 percent. Formal plans for doing so were required and progress monitored. Federal agencies in Atlanta (the region’s largest employer) also agreed to implement similar strategies. Finally, the EPD sought to

¹⁸ This case study was informed by document and literature review, including Henry (2002); Henry & Gordon (2001); Henry & Gordon, (2003); Henry & Rivera (1998). It was also informed by a 2002 interview with Gary Henry, one of the evaluation’s principal investigators and a Professor in the Andrew Young School of Policy Studies and the Department of Political Science at Georgia State University.

¹⁹ Ground-level ozone is an odorless and colorless gas undetectable to human senses. It decreases lung capacity in healthy adults, but more seriously affects children, seniors, and those with respiratory problems. Negative health effects can subside when ozone levels subside.

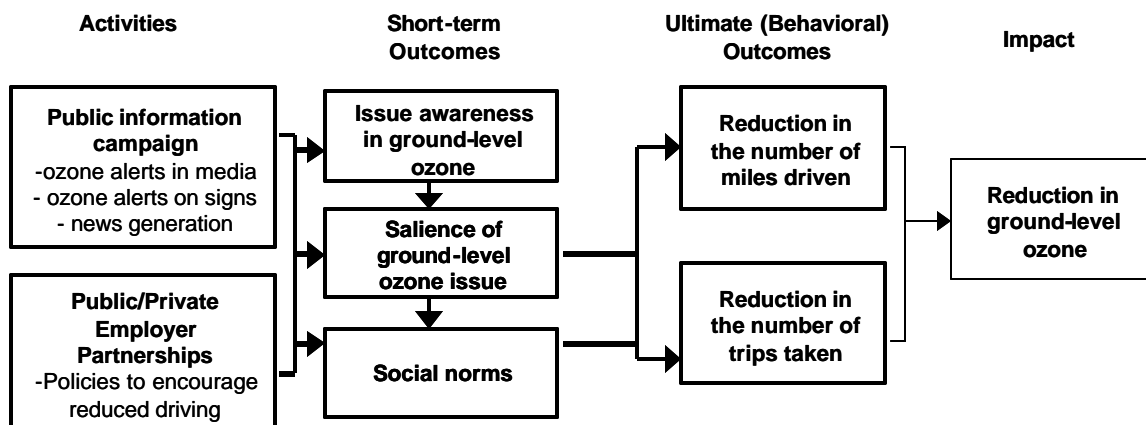
develop partnerships with private businesses to get their voluntary participation in the effort.

It is important to note that considerable front-end research, performed by the evaluators, went into the development of the VOAP public information campaign. While a discussion of front-end research is outside the scope of this paper, a few general points are important to mention here, as this research was both unique and informed the eventual campaign and evaluation design (see Henry & Rivera, 1998, for a more detailed explanation). First, the evaluator began with two related social science theories – the Theory of Reasoned Action (Ajzen & Fishbein, 1980) and Theory of Planned Behavior (Ajzen, 1985) – about how to effect individual behavior change and the variables that predict it. Using a survey conducted with a random sample of residents in the Atlanta region, the evaluator tested the theories' application to ozone-reducing behavior. Results helped to determine what behaviors the campaign should focus on (driving), and the factors that were most likely to affect those behaviors. For example, research revealed that many Atlanta residents were willing to drive less, but they felt as if they had little control over their behavior (e.g., because they had to be at work at a specific time). This finding helped to develop VOAP's push to try to create a social context among employers that supported the campaign's outcomes (for more information on this research see Henry & Rivera, 1998).

Theory of Change

Figure 6 illustrates the general VOAP theory of change. Activities included both components of the program – the public information campaign (ozone alerts, news stories) and the development of public and private employer partnerships.

Figure 6. Voluntary Ozone Action Program Theory of Change



The role of public/private partnerships in this model represents the belief that public and private employers who developed partnerships with the Environmental Protection Division and got behind the cause of reducing ground-level ozone, would ultimately make it easier for their employees to change their driving behaviors (e.g., by organizing carpools, allowing flexible work schedules, etc.).

VOAP's short-term outcomes included an increase in awareness of ground-level ozone issues, increased salience of the issue, and a change in social norms concerning what should be done about the region's ozone problem. Salience and social norms were then

expected to lead to behavior change in the form of reduced driving, both in terms of the number of miles driven and trips taken.²⁰ If less driving occurred, then presumably this would contribute to the reduction of ground-level ozone in metropolitan Atlanta.

Focus/Methods

Evaluators at the Applied Research Center in the Andrew Young School of Policy Studies at Georgia State University evaluated the campaign. The evaluation focused on the components that are bolded in Figure 6.

Rolling sample surveys (daily tracking surveys) were the primary assessment method. These surveys, which were innovative in this context, obtained measures from an independent sample of 32 Atlanta residents *each day*.²¹ Interviews took place between May 1, 1998 and September 31, 1998 – yielding a total of 2,935 responses over 153 days.²² Once respondents completed the survey, they could not be interviewed again. The independent samples were representative of the population in the 13-county Atlanta region (for more information on the use of rolling sample surveys, see Henry & Gordon, 2001).

The surveys, which were identical on alert and non-alert days, lasted an average of 7.1 minutes, and included 30 behavioral, awareness, and attitudinal items and 12 demographic items (Henry & Gordon, 2001). These included:

- Ratings of the importance of five issues from a personal and community standpoint (schools, jobs/economy, air quality, ground-level ozone, environment)
- Driving behavior over the last 24 hours (number of trips taken and miles driven)
- Mode of transportation to work in previous 24 hours (drive alone, mass transit, carpool, walk or bike, work at home, carpool)
- Awareness of ozone alerts
- Perceived efficacy (perceived ability to have an effect on ground-level ozone)
- Perceived personal health risks (poor air quality affects the person or the person's family directly).

Design/Analysis

The evaluation tested three core hypotheses:

- 1) The campaign affected the public's awareness of ground-level ozone.
- 2) Ozone alert days resulted in the reduction of miles driven and trips taken.
- 3) Ozone alert days resulted in reduced driving by government employees.

Analyses of the first hypothesis (using an ordinary least squares regression model) determined the variables that predicted awareness of ground-level ozone. The nine groups of variables tested as predictors included 1) awareness of ozone alert in past 24 hours, 2) demographics,²³ 3) day-of-week, 4) month, 5) commuting distance,²⁴ 6)

²⁰ Note that evaluators also explored a number of intermediate outcomes not illustrated in this model, such as avoiding driving during rush hours and limiting use of gasoline-powered lawn equipment.

²¹ Interviews were done during the late afternoon and evening hours, and were conducted every day except Memorial Day and the Fourth of July.

²² The response rate was 54%. Completed responses required an average of 6.7 call attempts. Attempts were made up to a maximum of 22 times.

²³ Sex, education, age, income, an income dummy variable for high-income, and minority status.

²⁴ Included two parts – one for residency and one for work location.

location of newspaper story,²⁵ 7) perceived community salience of ground-level ozone, 8) timing of alert,²⁶ and 9) attitudinal factors (perceived efficacy and health risks).

Analyses of the second hypothesis (using an analysis of covariance model) examined the variables that predicted driving behavior. The seven groups of variables tested included 1) awareness of ground-level ozone (resulting from the first analyses), 2) demographics, 3) day-of-week, 4) month, 5) commuting distance, 6) location of newspaper story, and 7) weather.²⁷

These analyses also examined if total miles or trips taken were reduced on alert days compared to non-alert days. Because ozone alerts happened on some days and not others, it was possible to develop a natural experiment with ozone alert days serving as the days when the treatment occurred, and non-alert days serving as the control days. Analyses compared behavior on alert days with behavior on non-alert days. This design allowed for a causal interpretation of whether the campaign caused changes in driving behavior.

The third hypothesis related to the question of whether voluntary involvement of business, government, or other employers made a difference in the campaign's behavioral results. As described earlier, most local, state, and federal government employers supported VOAP and took steps to encourage a social context in which it could be successful. Private employers were less proactive during the campaign's first year. As a result, the evaluation could compare government employees to non-government employees in terms of the extent to which their workplace made a difference in reducing miles driven and trips taken on alert days. If it found that only government employees reduced their driving, the evaluation could conclude that the campaign alone was insufficient to change driving behavior. It could *not*, however, conclude that the campaign was a necessary part of the causal package (because it was not possible to disentangle how the campaign, apart from employer behavior, would have fared). Analyses of this hypothesis looked at whether miles or trips taken were reduced for either government or non-government employees on alert days, and if miles or trips taken were reduced more for government employees than non-government employees on *either* alert or non-alert days.

Key Findings

About the campaign's effects on public awareness (1st hypothesis):

- Ozone alerts increased the amount of awareness about ground-level ozone. Those who were more aware of ozone said the issue was more important to them. Those personally affected by poor air quality were *not* more aware of ground-level ozone.
- Greater exposure to media messages affected awareness, as commuting patterns affected awareness (those who commuted during times when alerts were broadcast were more aware, as were those with longer commuting times).
- Articles on the front-page of newspapers increased awareness, but not articles on the Metro page.
- Those with higher incomes, older residents, and whites heard more about ground-level-ozone.

²⁵ Front page or metro section of the *Atlanta Journal Constitution*.

²⁶ If the alert was on the previous day or two days prior.

²⁷ Less driving was expected as the temperature or dew point rose, or if there was rain.

About whether ozone alert days resulted in driving behavior change (2nd hypothesis):

- Overall miles driven were significantly reduced on alert days (35.4 miles on non-alert days to 29.9 on alert days).
- Many predictors of driving behavior were as expected – men and individuals with higher incomes drove more, and those with the longest commutes drove more.
- Awareness did not have a significant effect on miles driven, though this did not mean that the campaign was not necessary for obtaining reductions in driving.

About whether ozone alert days resulted in reduced driving by government employees (3rd hypothesis):

- On alert days government employees reduced the number of miles driven and the number of trips taken by an average of almost one full trip, a statistically significant amount, while non-government employees did not.
- Changes in driving behavior were only significant when social contexts reinforced those changes (i.e., employers instituted policies to support the behavior change). Therefore it was concluded that campaigns can increase awareness, but changing well-established behaviors probably requires changes in the social context supporting that behavior change.

Evaluation Lessons

This was a resource-intensive and sophisticated evaluation. Its design and implementation required considerable evaluation and methodological expertise. The payoffs, however, were substantial in terms of obtaining definitive findings about the campaign's effectiveness and what it took to change driving behavior. This evaluation contributes much to the knowledge base about how to conduct evaluation in this field.

Rolling sample surveys are a useful evaluation method for this context. One of the most important lessons from this evaluation was methodological, through its use of rolling sample surveys. Adapted from political polling methods, this method, applied in this context, was extremely innovative and holds promise for other campaign evaluations, particularly when it is possible to “control” the intervention (e.g., knowing which days were ozone alert days). Other common survey methods create measurement and interpretation problems. For example, random sample surveys have biases and accuracy problems when it comes to remembering or predicting behavior, and repeated surveys with the same people over time have problems with recruitment biases and reactivity to repeated measurement. Where their use is possible, rolling sample surveys hold much promise for the field of campaign evaluation. They can aid in determining when the public gets worn out with a particular message, or how unexpected events during a campaign impact effects. They also allow for the testing of various communications-related theories (e.g., agenda setting) because of their continuous measurement of public opinion over time (Henry & Gordon, 2001).

Social science theory and front-end research informed campaign and evaluation planning. While again a discussion of front-end research is outside the scope of this paper, its use here was important to the evaluation. Here the evaluators used social science theory about known predictors of behavior change (Ajzen, 1985; Ajzen & Fishbein, 1980) to help determine how the campaign could be structured for maximum benefit. A pre-campaign survey of 1,275 Atlanta residents revealed a list of specific ground-level ozone reducing behaviors that Atlanta residents were more or less likely to do. Behaviors that were more likely included disposing of leaves without burning, cutting

back on miles driven, decreasing rush-hour driving, cutting back on car trips taken, and driving less often. Behaviors not likely to be changed included taking mass transit, waiting until the weekend to get gas, or telecommuting. The campaign then viewed the challenge in one of two ways – reinforcing behaviors that were likely or changing behaviors that were unlikely. It chose the former because the challenge with the latter was much greater. In turn, the evaluation then focused on driving behaviors (Henry & Rivera, 1998).

This early research also revealed other factors critical to the evaluation. First, in any campaign where behavior change is a goal, it is important that the evaluation measure *specific* behaviors. It would not have been useful here to ask Atlanta residents if they had “performed ground-level ozone reducing behaviors.” Specific behaviors had to be identified in both the campaign and the evaluation (i.e., driving less). In addition, front-end research revealed the potential importance of looking at a person’s *perceived control* over particular behaviors, as a moderator of whether those behaviors were adopted. While Atlanta residents expressed that they were not likely to telecommute, for example, it was unclear whether this was because they did not have the opportunity to do so, or whether they were choosing not to. This reinforced the point in the campaign, and the subsequent evaluation, that it was important to include and examine the effects of employer policies on behavior change (Henry & Rivera, 1998).

This evaluation answered questions about whether a campaign alone is sufficient to affect behavior. Finally it is important to recognize this evaluation for the substantive value of its findings. It determined that the information campaign alone was insufficient to change driving behavior. The larger lesson here is that while campaigns can increase awareness, changing well-established behaviors probably requires changes in the social context to support that behavior change.

D. Community Trials Project ²⁸

Purpose	Behavior Change	↔	Policy Change
Target Audience	General Public	Targeted Public	Policy makers
Evaluation	Experimental	Quasi-Experimental	Other/Non-Experimental

Research shows that availability of alcohol and high-risk alcohol consumption patterns (e.g., binge drinking, drinking before driving, and underage drinking) can increase a community’s risk of alcohol-involved traffic accidents and violent assaults. The Community Trials Project tested the potential of a five-year (1992-1996) comprehensive community prevention strategy in three communities (of approximately 100,000 residents and located in Northern California, Southern California, and South Carolina) to

reduce the incidence of such alcohol-related risk factors and outcomes.

The project was based on a scientific-community partnership, in which the research team identified the core elements of the overall prevention strategy (see below) and

²⁸ This case study was informed by a structured telephone interview with Andrew Treno, Research Scientist, at the Prevention Research Center. It was also informed by document and literature review, including Holder, Gruenewald, Ponicki, Treno, Grube, Saltz, Voas, Reynolds, Davis, Sanchez, Guamont, & Roeper (2000); Holder & Treno (1997); Treno, Breed, Holder, Roeper, Thomas, & Gruenewald (1996); Treno & Holder (1997a); and Treno & Holder (1997b).

provided training, technical assistance, and resources around those elements. The project teams in the experimental communities then tailored the strategy and prevention elements to their own needs and circumstances (Treno & Holder, 1997a).

The research-based prevention strategy had five interrelated elements. At the time the project started, the efficacy of each element alone had already been proven; however they had never been used together to reduce alcohol-related injuries and death (Holder, 1993; Holder et al., 2000; Treno, Breed, Holder, Roeper, Thomas, & Gruenewald, 1996):

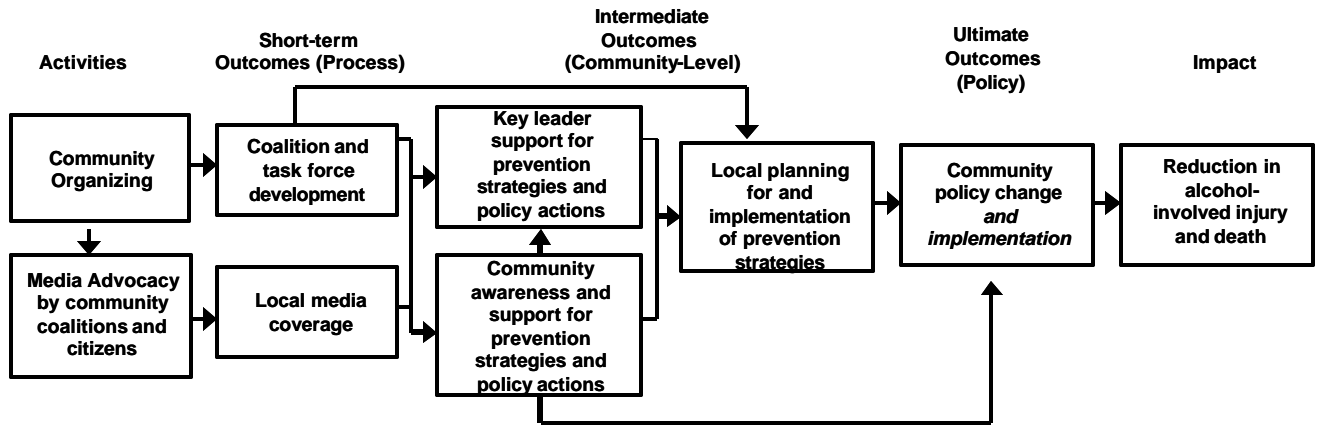
- 1) *Community knowledge, values, and mobilization* – This component integrated two main aspects – community organizing and media advocacy, and was used to support the other four prevention elements that follow. It included the development of community coalitions and task forces to implement specific interventions and to increase community concern about alcohol-involved trauma (alcohol-involved traffic crashes that led to injuries and fatalities). Coalitions also implemented media advocacy activities to build community support for the four policy-level prevention elements that follow.
- 2) *Responsible beverage service* – This component was designed to reduce the probability that patrons leaving licensed alcohol serving outlets would be intoxicated or underage. It included incentives for bars and restaurants to provide server and staff training to identify intoxicated and/or underage customers in bars and restaurants, and to strengthen beverage service policies to prevent customer intoxication and driving under the influence.
- 3) *Underage drinking* – This component was designed to reduce the social and retail availability of alcohol to underage persons. Included were school-based and community programs for parents and kids about sales and access to alcohol by minors, and the training of off-premise retailers to reduce sales to minors. It also promoted the enforcement of underage sales laws.
- 4) *Drinking and Driving* – Intended to increase the actual and perceived risk of being apprehended while driving under the influence (DUI) of alcohol, this component included media advocacy to increase community support for and awareness of DUI detection. It also included greater law enforcement efficiency and effectiveness in detection, which incorporated increased DUI checkpoints, training for police officers in new techniques for identifying drunk drivers, and the use of passive alcohol sensors to increase the probability of positive detection.
- 5) *Access to Alcohol* -This component included the use of local zoning policy to control alcohol outlet density in order to reduce the retail availability of alcohol in communities.

The focus of this case study is on the first element of the five-part prevention strategy, community mobilization, which was considered a critical support for the implementation of the other four elements. Community mobilization is defined here as “the purposeful organization of community members to implement and support policies that will reduce alcohol-involved strategies” (Treno & Holder, 1997a, p. S175). Two main components constituted mobilization in the Community Trials Project – community organizing and media advocacy. Ultimately, mobilization’s objective was the implementation of specific community-level policies, connected to the prevention elements above, which had the potential of reducing alcohol-involved injuries and death (Treno & Holder, 1997a).

Theory of Change

Figure 7 shows the theory-of-change behind community mobilization in the Community Trials Project (developed from conceptual models in Treno et al., 1996; Treno & Holder, 1997a; and Treno & Holder, 1997b).

Figure 7. Community Trials Project Community Mobilization Theory of Change



The overall theory of change represents the Project's assertion that community mobilization alone (with the components of community organizing and media advocacy), cannot impact alcohol-involved problems directly, but that they should be used in concert with and support the other environmental prevention elements (the remaining four elements of the strategy defined above). In this way mobilization is seen as a means of supporting local prevention elements that lead to policy change and implementation, which in turn can decrease alcohol-involved injury and death (Holder & Treno, 1997).

The first short-term outcome of community organizing was the development of local coalitions in the three communities to work on alcohol-involved issues. Those coalitions then formed task forces around the four remaining elements of the prevention strategy (i.e., responsible beverage service, underage drinking, etc.). The coalition was expected to increase community awareness and support in the community for these strategy elements, while the task forces simultaneously increased key leader support.²⁹ Community awareness and support were also expected also to have an influence on key leader support.

Media advocacy, implemented by both coalitions and community members, was designed to use the local media to reshape news content in support of prevention strategy elements. The idea was that local news media would raise awareness of a specific alcohol problem, spotlight solutions to alcohol problems, and put pressure on key leaders to adopt those solutions (Holder & Treno, 1997).

The link between media advocacy and policy change is based on the idea that media coverage will stimulate and support the adoption of relevant alcohol policy by 1) directly increasing key leader support for the specific environmental interventions being

²⁹ Key leaders are defined as decision makers who can actually bring about policy changes, including city or county council members, public agency representatives (e.g. heads of police departments or school superintendents), and private business managers (e.g. owners and operators of licensed alcohol establishment) (Holder & Treno, 1997).

promoted, and 2) increasing community awareness and mobilizing public action and support, which is in turn expected to have an effect on key leaders and consequently policy change and implementation. “Even if local decision makers do not support and are even uninterested in an issue, effective media advocacy can change local priorities if the decision maker recognizes the level of community importance assigned to a specific alcohol problem and the associated preventative policy” (Holder & Treno, 1997, p. S190).

Focus/Methods

The Prevention Research Center (PRC) of the Pacific Institute for Research and Evaluation in Berkeley, California evaluated the Community Trials Project. PRC is a national center for prevention research, with a focus on understanding the social and physical environments that influence the individual behavior that leads to alcohol and drug misuse.

A main objective in this project was research and learning about the effectiveness of the prevention approaches used, therefore this project was also known as a prevention trial. As such, it used three “experimental” communities where the intervention took place, matched with three “comparison” communities where the intervention did not, allowing for an examination of differences of effects between the experimental and comparison communities.

This evaluation is one of very few that has attempted to systematically assess the links between the types of components depicted in Figure 7. It attempted to determine whether the intervention as planned was actually implemented, and whether the theory of change worked as intended.

Structured weekly project reports and event and intervention forms completed by local project coordinators were the primary data sources for the community mobilization elements in Figure 7. These reports and forms captured:

Community organizing activities – e.g., recruitment and hiring, staff training and technical assistance

Media advocacy activities – e.g., press conferences, events to which media are invited, editorial writing

Coalition and task force development – e.g., establishment of members and coalition bylaws, completion of media advocacy training

Key leader support for prevention strategies and policy change – e.g., police participation in media events, city council endorsement for various projects, police chief support for increased DUI enforcement

Local planning for the implementation of prevention strategies – e.g., special DUI grant developed, planning developed for change in city zoning of alcohol outlets

Community policy change and implementation – e.g., city council adoption of mandatory server training (policy), city establishment of zoning for alcohol outlets (policy), police use of special DUI enforcement equipment (implementation).

Evaluators verified the information in these reports in various ways. For example, in the case of media advocacy activities, technical consultants validated this information.

These data were then entered into an electronic database and used to compile *chronologies* or timelines of critical activities in each community to determine the sequencing, timing, type, and extent of critical events. Local project coordinators were interviewed to verify this information (Treno & Holder, 1997a).

Ongoing content analysis of local newspaper coverage of alcohol-involved issues in both the experimental and comparison communities, and local television coverage in the experimental communities, assessed *media coverage and content*. Newspaper articles from daily and weekly newspapers in the six communities,³⁰ and published over four years of the project's implementation, were coded using a structured protocol. Articles were coded for date, placement, size, type (e.g., editorial, feature), geographic area discussed, and subject matter conforming to the elements in the prevention strategy (e.g., responsible beverage service, alcohol outlets, etc.). Coding also included control topics (e.g., drug abuse and enforcement).

To aid in analysis and come up with a more meaningful measure of media coverage, the evaluators developed a *composite news score* that compiled (Holder & Treno, 1997a; Treno et al., 1996):

- Total number of stories
- Total area or time allotted to each story
- Total number of news stories above average length (18 column inches)
- Total stories with pictures and graphics
- Total stories on the front page or in the local TV news program

Composite scores were calculated by month and by community and location covered, resulting in a total of 270 separate data points (6 communities across 45 months).

Community awareness and support – These outcomes were assessed using monthly community telephone surveys (for 66 months) of 120 randomly selected individuals in each of the experimental and comparison communities. The survey included many items that measured variables such as media exposure, awareness of alcohol-involved problems, support for specific prevention strategies, and perceived risk of arrest (Treno & Holder, 1997b).

Reduction in alcohol-involved injuries and death was tracked with six self-report items about alcohol consumption and driving on the community survey, monthly longitudinal traffic record data on car crashes, and rates of alcohol-related crashes and assault injuries observed in emergency departments and admitted to hospitals.

Design/Analysis

The evaluation used a longitudinal multiple time series design, and compared data from experimental and matched comparison communities.

The evaluation of media advocacy involved determining whether 1) media activities and events resulted in increased media coverage in experimental communities compared to comparison communities, and 2) whether media coverage led to its intended effects. The examination of those questions involved multiple types of analysis. For the first question, for example, evaluators developed a time series plot of the composite news

³⁰ Articles were selected from the first two sections of each daily newspaper, and the editorials and letters to the editor. Articles were selected from the entire weekly newspapers.

scores described above for experimental and comparison communities (aggregated for each). They then mapped onto that plot the different stages of project activity drawn from the community chronologies (pre-implementation, media campaign, ongoing media advocacy, etc.). Numerous statistical analyses were computed to determine whether significant differences existed in news coverage over time resulting from media advocacy activities in the experimental communities, compared to the comparison communities.

The media advocacy evaluation also examined the connection between media coverage and community awareness about various problems and proposed policies. Analyses plotted together and examined over time the two measures to determine their relationship in both the experimental and comparison communities.

The longitudinal chronologies (which were 20 pages and more) were analyzed qualitatively to determine the chronological concurrence of the community mobilization process to its community-level outcomes, including policy changes and implementation. In other words, evaluators looked to see if there was a logical sequence within each community from community activities and process, to key leader support for the prevention strategies and policies being promoted, to the adoption of those policies and their implementation.

Key Findings

About media advocacy's effects in producing media coverage:

- Media advocacy was successful in increasing both print and television news coverage of local alcohol-related topics in the experimental communities. A statistically significant difference existed between news coverage resulting from media advocacy in experimental communities compared to comparison communities (Holder & Treno, 1997).

About the link between media coverage and key leader attention:

- Increased media coverage can focus key leader attention on problems and prevention policies or solutions (Holder & Treno, 1997).
- Successful media advocacy to achieve policy change requires specific policy objectives and the awareness and support of those policy objectives by key leaders and decision makers. News attention is insufficient without specific policy goals and community organization to support those goals (Holder & Treno, 1997).

About the effects of the prevention strategy elements and policy change and implementation:

- Community mobilization achieved its ultimate outcome (local implementation of policy changes) in all three experimental communities (Treno & Holder, 1997a).
- Evidence of policy change and implementation related to the components of the five prevention elements included (Holder et al., 2000):
 - o Responsible beverage service – policies at bars and restaurants changed; 409 managers and servers were trained.
 - o Underage drinking – underage alcohol purchases reduced from 44% to 17%.
 - o Drinking and driving – 410 sobriety checkpoints were established.
 - o Access to alcohol – stricter zoning requirements were implemented in two of the three communities.

About the intervention overall:

- The comprehensive community-based intervention reduced high-risk alcohol consumption and alcohol-involved injuries resulting from traffic crashes and assaults.
- Comparing experimental and comparison communities (Holder et al., 2000):
 - o Self-reported rates decreased: amount of alcohol consumed per drinking occasion declined 6%; having too much to drink declined 49%; driving over the legal limit decreased 51%.
 - o Traffic data showed that nighttime injury crashes declined by 10% and crashes in which the driver had been drinking declined by 6%.
 - o Assault injuries in emergency departments declined by 43%; all hospitalized assault injuries declined by 2%.

Evaluation Lessons

This was a resource- and time-intensive evaluation. Its yield, however, in terms of substantive findings and methodological value, is great.

This approach is instructive about the methodology of evaluating community organizing, media advocacy, and community mobilization. While numerous case studies of media advocacy have been done and are useful for learning (e.g., Gerber & Dorfman, 1996), neither media advocacy nor community organizing, or their effects on policy change, is typically systematically evaluated. This evaluation was methodologically important for the field of communications evaluation. For example, evaluators:

- Used a *composite media measure* for media analysis – Much analysis of media coverage merely provides basic information on article counts. The composite measure approach used here took into account other important factors in the measurement of media advocacy efforts and effects, such as article size, placement, and the presence of graphics.
- Used event and intervention forms to develop *chronologies* – This unique approach provided a method for systematically assessing the chronological concurrence of mobilization activities, with key community leader support, and policy change and implementation. Because this approach was implemented after the evaluation had started, the chronologies were time- and labor-intensive to develop. With reporting formats that coincide with evaluation data needs, however, this approach is more efficient and holds promise for the many other evaluations that face this type of documentation challenge.

This evaluation contributed important substantive findings about community organizing and media advocacy. We know relatively little about the causal effects of community organizing and media advocacy on policy change. This evaluation tested this causal chain. While the evaluation could not separate out effects to determine if either media advocacy or community organizing *alone* could produce policy change, it found that media advocacy in combination with community organizing, had an effect.

While not discussed earlier in the case study, another important finding was about the relative utility of public information campaigns compared to media advocacy. The evaluation provided the opportunity to compare and contrast public information campaigns and media advocacy in one experimental and matched comparison community. It found that media advocacy increased community awareness independent

of planned information campaigns – a finding that if replicated in other situations could have major cost and resource implications (Holder & Treno, 1997).

E. The Devolution Initiative ³¹

Purpose	Behavior Change	↔	Policy Change
Target Audience	General Public	Targeted Public	Policy makers
Evaluation	Experimental	Quasi-Experimental	Other/Theory of Change

During the mid-1990s, several major policies passed down or devolved decision-making authority and responsibility for key policy issues affecting low-income children and families from the federal level to the state and local levels. Paramount among these was passage of the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) of 1996, which gave states the authority to make their own decisions about welfare reform policies.

In response to this trend on both welfare reform and health care issues (e.g., the State Child Health Insurance Program), the W.K. Kellogg Foundation formed the Devolution Initiative, which from 1996 through 2001, supported 30 national and state research, policy, and advocacy organizations, and teams of minority researchers and community organizers, to work together – with a particular focus in five states (Florida, Mississippi, New York, Washington, and Wisconsin) – on three primary goals:

- To create an objective information base about the impact of welfare reform and health care devolution that is useful and useable to a broad group of stakeholders, including community members
- To share the findings with policymakers and the public
- To use the information and other community resources to promote public participation in informing policy agendas and decisions.

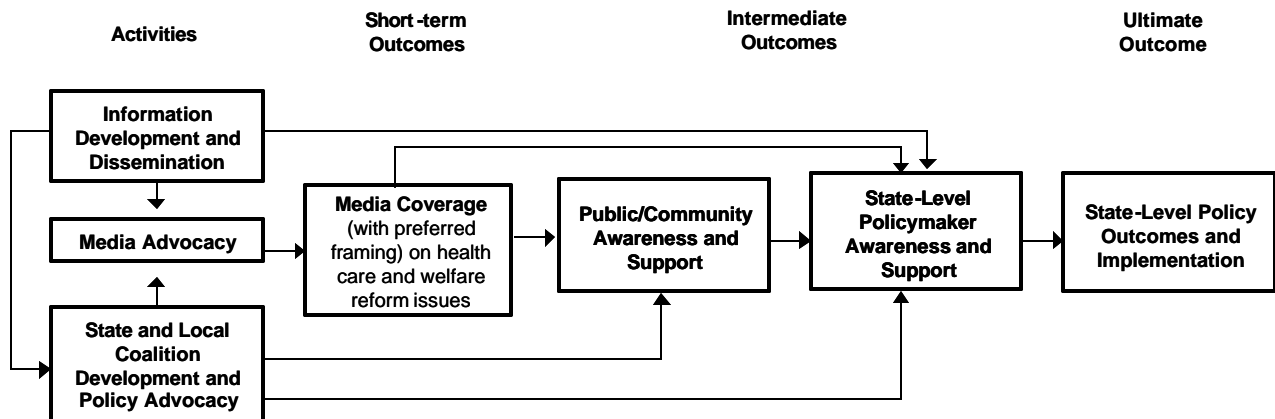
Initiative participants used a number of mutually-supporting activities to achieve these goals. All participants (which included both national and state organizations) engaged in information development (e.g., research, evaluation, and policy analysis) and dissemination. The Initiative also featured an investment in statewide advocacy organizations in the five states to build coalitions, engage in community mobilization, and build the capacity of those affected by devolving welfare reform and health care policies to participate in and inform the policy process. The Initiative ultimately strove to inform the adoption and implementation of devolution-related policies in states.

Theory of Change

Figure 8 presents the theory of change behind the Devolution Initiative.

³¹ This case study was largely informed by direct experience, as the author participated in the evaluation of the Devolution Initiative. Other sources included Harvard Family Research Project (2001, 2002).

Figure 8. Devolution Initiative Theory of Change



Information development and dissemination played a prominent role. As welfare reform and health care decision making devolved to the state and local levels, the belief was that policymakers who found themselves faced with decisions about new devolution-related issues would need and want to better understand the issues in order to make good and informed decisions. Consequently, the Initiative invested substantially in research, monitoring, and policy analysis to track what was happening as a result of devolution, and to disseminate that information, in diverse and user-friendly formats, in states. Initiative participants developing information were expected to reach a broad range of audiences, including policymakers directly, and advocates at the state and local levels who would then use that information in their policy advocacy activities.

The Initiative theorized that media advocacy would be another way to get the messages and implications coming out of research, monitoring, and policy analysis to this wide array of audiences. All of the Initiative's participating organizations (national and state) conducted media outreach and advocacy for the purpose of getting media coverage and visibility for the issues and positions they focused on. The hope was that this media coverage would have two main effects – to increase public awareness and build the support necessary to place pressure on policymakers to make good and informed policy decisions, and to increase policymaker (state-level specifically because of devolving authority) awareness and support directly. Media advocacy activities most often took the form of events (e.g., press conferences or events that attracted media attention) that coincided with “policy windows” of influence (Kingdon, 1984), relationship building with reporters and editorial boards, and the dissemination of press releases, op-eds, and editorials to highlight and reframe issues.

For the Devolution Initiative participants, their concern was not only increasing the quantity of coverage and the visibility of their messages in it, but the quality of it as well. As such, a considerable amount of front-end effort was spent in understanding how the media had been framing welfare reform and health care stories. They learned that in the mid-1990s, as welfare reform began, media coverage historically had tended to place responsibility on low-income individuals for their plot in life, and gave those same individuals responsibility for pulling themselves out of poverty (Gamson & Lasch, 1983; Heintz-Knowles, 1998; Iyengar, 1991; Ryan, 1996). Initiative participants felt that these frames were counterproductive to their messaging, and that the social and institutional

reasons for poverty needed to be highlighted more, along with the government's role in helping people to rise out of poverty.

Within the five focus states, other activities complemented and supported information development and media advocacy – namely statewide and community-based coalition development and policy advocacy. Coalition activities – such as statewide meetings, rallies at the capitol, one-on-one meetings with policymakers, and other activities or events – were implemented either to mobilize community awareness and support, or to inform policymakers directly.

Focus/Methods

The Harvard Family Research Project (HFRP) based in Cambridge, Massachusetts evaluated the Devolution Initiative. The evaluation assessed the theory of change components that are bolded in Figure 8.

A systematic document review assessed *information development*. Four years of participant documents and materials were collected and analyzed to track indicators about their content, format, and target audiences.

Dissemination, media advocacy, and coalition development and policy advocacy activities were tracked using multiple methods that included a combination of regular key informant interviews and questionnaires, site visits, participant observation, and media tracking (see below). For media advocacy activities, the document review described above and other qualitative assessment determined the types of messages that Initiative participants disseminated.

The evaluation used a method labeled media tracking to examine *media coverage*. This process began with the identification of news sources related to the media advocacy work of Initiative participants. Because Initiative participants conducted media outreach at both the national and state levels, the 29 electronic and print sources chosen included five newspapers,³² and four broadcast services,³³ with a national reach; the two newspapers with the largest circulation in each of the five focus states; and ten ethnic media sources with either a national circulation or readership in one of the five states (Initiative media advocacy included targeted outreach to the ethnic media).

Using carefully-selected search terms and criteria, evaluators searched for welfare reform and health-care-related articles using electronic news searching services (e.g., Lexis-Nexis), or by reading non-electronically-available newspapers in their entirety. Once the population of welfare reform and health care-related articles was identified, a quarterly sample was selected for analysis. A total of 2,391 articles – 552 on welfare reform, and 1,839 on health care – were selected. Content analysis followed a structured coding protocol, and assessed six main aspects of the coverage: 1) amount of overall coverage, 2) issues covered, 3) frames used in coverage, 3) messages conveyed, 4) mention of groups affected by devolution, and 5) sources used to inform coverage.

The evaluation documented *public/community awareness and support* at the state level by tracking various state and community organizing and mobilization indicators (e.g.

³² USA Today, Wall Street Journal, New York Times, Los Angeles Times, Washington Post

³³ ABC News, NBC News, CBS News, NPR

coalition membership over time, participation levels in policy advocacy) using key informant interviews and surveys.

Questionnaires and interviews with state-level policymakers, staffers, and administrators responsible for decision making about health care and welfare reform issues assessed indicators of *policymaker awareness and support*.

Policy outcomes and implementation related to the policy focus areas in each state (e.g., immigrant access to health care; child health insurance eligibility and enrollment; child care access, cost, and quality) were tracked using a tailored policy tracking system in each state, which relied on regular searching of a set of legislative, media, and policy alert electronic sources. Regular interviews with Initiative participants verified this information.

Design/Analysis

The evaluation design followed a “theory of change” approach. This approach requires the articulation of the theory of change at the beginning of the Initiative, which then serves as a guide for the Initiative’s ongoing strategic development and management. Evaluators then seek evidence that the components of the theory are in place and that the theorized links between the activities and their intended outcomes exist (Connell, Kubisch, Schorr, & Weiss, 1995).

The Devolution Initiative evaluation did not attempt to determine whether the components of the theory of change were *causally* linked in the five focus states. Rather, the focus was on the independent assessment of each component, using multiple methods where possible, followed by a longitudinal case study approach in each of the five states that examined qualitatively and sequentially the relationships between the components.

The primary question was – Can a *plausible and defensible* case be made that the theory of change worked as anticipated and that Devolution Initiative investments had their intended effects? For example, for media advocacy, the evaluation combined the stand-alone quantitative analysis of the media coverage, with a qualitative assessment of whether that coverage matched the intent of the messages Initiative participants were sending through their media advocacy activities.

Key Findings

About information development and dissemination:

- Over 1,200 documents resulted from Initiative investments in research, monitoring, and policy analysis. Information was targeted to an appropriately diverse range of audiences and was user-friendly in format.

About media advocacy and coalition development and policy advocacy activities:

- The Initiative supported nine statewide coalitions in the five states with memberships that grew collectively to over 800 individuals and organizations. All coalitions increased their membership under the DI and implemented media advocacy and policy informing activities.

About media coverage:

- *Amount:* While articles about health care coverage outnumbered articles about welfare reform by more than three to one, coverage overall in terms of quantity was limited for both issues.
- *Framing:* There was more evidence of the social responsibility frame in media coverage of both health care and welfare reform than the individual responsibility frame – a finding that resonated with Initiative participants' intents.
- *Sources:* Researchers, policy experts, and advocates – the types of organizations supported under the Devolution Initiative – were among the top five sources cited in both health care and welfare reform articles. Private corporations and federal policymakers were top sources in health care. Public administrators and policymakers were most often cited in welfare reform articles. While Initiative participants worked to change this, the voices of those affected by devolution were rarely featured or used as sources.

About public/community awareness and support:

- Each coalition's membership grew and diversified under the Initiative, with a focus on adding nontraditional partners and underrepresented groups.
- The Initiative mobilized and nurtured a large number of new advocates (though few new leaders emerged). These new voices were mobilized to inform local, state, and federal policy through mechanisms such as testimony at legislative hearings, one-on-one meetings with policymakers, and participation in policy advisory groups.

About policymaker awareness and support:

- State policymakers reported using Initiative participants and information as sources to inform their policy decisions, though the findings about this component in the theory of change were somewhat inconclusive due to measurement issues and challenges.

About state-level policy outcomes and implementation:

- A plausible case existed that the theory of change worked as intended in the five states and that state-level advocates and coalitions informed 22 state-level policy issues in 1999-2000 and 47 issues in 2000-2001.
- Intended policy outcomes, including results in which advocates successfully "held the line" (e.g. avoided cuts to funding or eligibility), were achieved for 17 of the 22 issues grantees informed in 1999-2000 and 32 of 47 issues in 2000-2001.

Evaluation Lessons

This evaluation was both time- and resource-intensive, covering six years and requiring a large evaluation investment. In essence, the evaluation used multiple methods to assess each component in the theory of change, and then assessed qualitatively whether a plausible case could be made that those components were linked as expected.

Theory of change evaluation has value for policy change campaigns. The theory of change approach to evaluation, which was used with the Devolution Initiative, is growing in popularity among evaluations of complex social initiatives that emerge and change over time and try to effect change on multiple levels. It is a response to traditional evaluation approaches that cannot adequately deal with this type of complexity and lack of control over the intervention. Since most policy change campaigns share this type of

complexity, this approach could have much value in this field. While this case study featured a large-scale evaluation, it is possible to adapt and implement this approach on a smaller scale. In addition, while this evaluation attempted to provide a plausible case, and not a conclusive one, about relationships between the theory of change components, it is possible to incorporate methodological elements into this design, such as comparison groups or inclusion of counterfactuals, that can strengthen confidence in interpretations about effects and causation (e.g., Weitzman, Silver, & Dillman, 2002).

There is value in looking at multiple factors in the analysis of media coverage. Drawing from the experience of others who have raised the importance of including media frames and other variables about media *content* in research related media advocacy (e.g., Bales, 2002; Iyengar & Kinder, 1987), this evaluation attempted to include an assessment of frames, issues covered, messages, and article sources in its analysis of welfare reform and health care media coverage. There were challenges to this approach, including that it can be extremely resource-intensive, difficult to identify the right set of codes and ensure inter-coder reliability when coding is subjective. In addition, while there was considerable formative value in this approach for Initiative participants, questions exist about whether the strength and length of the intervention was strong enough to produce desired changes, particularly regarding framing. Many of these issues will require more attention in future evaluations of media advocacy, but the fact remains that when it comes to media advocacy, which tries to affect both quantity and quality of media coverage, the further search for efficient ways of incorporating quality into evaluation efforts is needed.

Campaign information needs and stage of development should factor into evaluation design. It is important to note here that the Devolution Initiative evaluation focused heavily on learning and continuous feedback to inform the Initiative as it developed. This fact influenced greatly how the evaluation was designed. For example, it was a widely accepted premise in the theory of change that media advocacy was an important part of trying to achieve policy change. Therefore questions about causal effects were somewhat less important than understanding how the media was covering the issues Initiative participants were working on, and then feeding that information about needs and opportunities back into the Initiative. This required the analysis of coverage content described above. Because this in turn required considerable resources to achieve, the tradeoff was between focusing resources on the systematic analysis of media coverage and feeding back information regularly, and developing an evaluation design that would assess causal effects but might not return results until the end of the initiative. The evaluators chose the former.

V. CROSS-CASE STUDY LESSONS

Much has been said already about the fact that the field of communications evaluation has a long way to go in its development. There are countless challenges that make evaluation difficult and progress incremental. But the field is at the point where it needs to stop using the challenges as excuses, and address more clearly *how* to tackle those challenges. The examples in this paper were designed to illustrate different responses to the “how” question. Characteristics of the evaluations featured that should be more prevalent in this field include:

- ? Evaluation based on a sound (and where possible research-based) conceptual model of how the campaign will achieve social change (e.g. theory of change)
- ? Front-end research that informs where the campaign *and* the evaluation should focus (too often it is only the former or neither).
- ? Recognition *and acceptance* that different evaluation needs and capabilities require different evaluation designs (and that causation is not always the most important question). But at the same, recognition that leverage to convince sponsors to invest in campaigns will be enhanced by causal analysis of campaign impacts.
- ? Recognition of the interpretive boundaries associated with specific evaluation designs (and then staying within those boundaries).
- ? A focus on evaluation for learning and continuous improvement (often related to a collaborative relationship between the evaluation and campaign staff).
- ? A commitment by campaign staff to be resourceful when necessary and find ways to track data on their own when possible.
- ? A push for methodological innovation and rigor when possible, and not always relying on the same approach (i.e. random sample or repeated sample pre- and post-campaign surveys).
- ? A recognition that evaluation *can* respond to the hard-to-answer questions (e.g. whether attention to social context is a necessary ingredient of behavior change, or whether media advocacy can contribute to policy change).

Below is a set of overarching case study lessons that expand on these factors, and have implications for how future work to build evaluation capacity can proceed. Lessons are arranged in the same organizing categories used in the case studies.

Theories of Change

Sound (and where possible research-based) theories of change help to avoid ineffectual evaluation. As the case studies illustrate, while there are some similarities to the way campaigns are designed, each is unique. Their focus in the short- and longer-term depends on many factors, such as how new the issue is, who it is targeting, and the type of change being sought. An alarming number of evaluations, however, discount the fact that campaigns are different, and use cookie cutter evaluation approaches. Such generic approaches are not as common in other fields. There could be a number of reasons why it happens in this one, including that

the theory behind how communications efforts achieve their intended change is largely unknown to many or has not been thought through, or the fact that the tools for facilitating the process of thinking through what campaigns are trying to accomplish in the short- and longer-term are not available, or widely used. This paper tried to make some progress on each of these accounts, by using one well-known evaluation tool – the theory of change. Future work needs to advance both the availability of tools like this for evaluation planning, and the sophistication of our collective thinking about how theories of change look (and work) for different types of campaigns.

Focus/Methods

Measurement challenges often require innovative methods. Target audience surveys done before and after a campaign will probably always be a standard method in this field. The case studies revealed, however, that other innovative methods exist for some challenging evaluation problems. Approaches like the rolling sample surveys used with the Voluntary Ozone Action Program have much more potential (and interpretive power) than traditional surveys if the conditions are right. The activity and event chronologies in the Community Trials Project, which were used to link community organizing and media advocacy to policy change and implementation, can be extremely useful for policy change evaluations and do not require extensive methodological expertise to create. Media tracking, as used in the Community Trials Project and Devolution Initiative, can easily yield much more information than the typical article counts. And the unique “in-home methodology” used in the Safe Gun Storage campaign to assess brand retention and ad recall is another example of how methods can be innovative in this context. More methods like these exist and need to be shared; still more are waiting to be developed.

Design/Analysis

Along with purpose, the campaign’s stage of development and information needs should factor into evaluation design. The case studies here represented a diverse array of evaluation approaches, including experimental, quasi-experimental, and non-experimental. The intent here was not to say that one approach is better than other; it was to say that different questions require different approaches. The Safe Gun Storage Campaign evaluation design - repeated random sample surveys – provided the opportunity to judge effects, with some interpretive constraints, but also with a reasonable time and resource investment. In the case of STOP IT NOW! VERMONT, the primary need was to test some of the assumptions in the emerging theory of change and to collect data that would be useful for learning. The design did not set out to prove if the campaign caused effects observed, but its designers were not trying to answer that question. They wanted to collect useful data that would help the campaign to know if they were on track and where to go next. In the case of the Voluntary Ozone Action Program, federal funding was at risk if ground-level ozone levels were not reduced. The campaign needed to determine and demonstrate if the investments they made worked. Consequently, a response to that question required an experimental design. With the Community Trials Project, one of the main purposes was to find prevention strategies that worked and could be replicated, therefore it used a quasi-experimental approach. The Devolution Initiative’s theory of change evaluation was designed to both fit with an evolving and complex policy change initiative and promote learning, therefore it examined

intended or theorized results compared to actual results at regular intervals throughout the six-year effort.

While the introduction to this paper stated that there is a need for more methodological rigor in general in this field, this does not mean that every evaluation should use an experimental or quasi-experimental approach. This often is not even feasible since these types of designs typically require much time and money. Rather than promote one design over another, the better solution is to help people make appropriate and acceptable choices based on the campaign's needs and what resources are available.

Key Findings

Evidence is growing about the conditions necessary for success. In addition to offering practical information about different ways of approaching evaluation, this set of case studies offers some substantive findings of interest that contribute to the growth of our knowledge base about what it takes for campaigns to achieve social change. Two of the case studies featured in particular – the Volunteer Ozone Action Program and the Community Trials Project – offer some conclusive findings about key questions that surface repeatedly in this field. They serve as guidance, or evidence if you will, about the conditions that are essential for effectiveness for behavior and policy change. The Volunteer Ozone Action Program evaluation told us that information alone was insufficient to produce behavior change; attention to social context was necessary. While this is not a new supposition, this evaluation confirmed it. The Community Trials Project evaluation showed us that community organizing combined with media advocacy can lead to effects on media coverage, key leader support, and subsequently policy change. For policy change campaigns, such evidence based on systematic evaluation is rare, and important. Cumulative evidence from these evaluations and others are essential to the continuing development of knowledge about how to design and evaluate campaigns that have better chances of success.

REFERENCES

- Advertising Council (in press). Case study of safe gun storage campaign (Draft). New York: Author.
- Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhi & J. Beckmann (Eds.), *Action-control: From cognition to behavior* (pp. 11-39). Heidelberg: Springer.
- Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Upper Saddle River, NJ: Prentice Hall.
- Bales, S. (2002). *The FrameWorks perspective: Strategic frame analysis*. Retrieved May 1, 2002, from <http://www.frameworksinstitute.org/strategicanalysis/perspective.shtml>.
- Bohan-Baker, M. (2001a). A conversation with Susan Nall Bales. *The Evaluation Exchange*, 7(1), 10-11.
- Bohan-Baker, M. (2001b). Pitching policy change. *Evaluation Exchange*, 7(1), 3-4.
- Children's Defense Fund (2002, September). *Protect children instead of guns: Fact sheet*. Retrieved February 15, 2003 from http://www.childrensdefense.org/ss_gvfs_protchild.php
- Coffman, J. (2002a). Beyond the usual suspects. *The Evaluation Exchange*, 8(3), 12.
- Coffman, J. (2002b, April). *Public communication campaign evaluation: An environmental scan of challenges, criticisms, practice, and opportunities*. Cambridge, MA: Harvard Family Research Project.
- Connell, J.P., Kubisch, A.C., Schorr, L.B., & Weiss, C.H. (1995). *New approaches to evaluating community initiatives: Concepts, methods, and contexts*. Washington, D.C: The Aspen Institute.
- Cook, T.D., & Campbell, D.T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin.
- Cook, T.D., Campbell, D.T., & Peracchio, L. (1990). Quasi-experimentation. In M.D. Dunnette and L.M. Hough (Eds.), *Handbook of Industrial Organizational Psychology* (pp. 492-576). Palo Alto, CA: Consulting Psychologists Press.
- Dorfman, L., Ervice J., & Woodruff, K. (2002, November). *Voices for change: A taxonomy of public communications campaigns and their evaluation challenges*. Berkeley, CA: Berkeley Media Studies Group.
- Dungan-Seaver, D. (1999). *Mass media initiatives and children's issues: An analysis of research and recent experiences relevant to potential philanthropic funders* [Electronic version]. Paper prepared for the McKnight Foundation.

- Fishbein, M., Triandis, H.C., Kanfer, F.H., Becker, M., Middlestadt, S.E., & Eichler, A. (2001). Factors influencing behavior and behavior change [Electronic version]. In A. Baum, T. Revenson, & J. Singer (Eds.), *Handbook of Health Psychology* (pp. 3-16). Mahway, NJ: Lawrence Erlbaum.
- Frumkin, P. (2002). *Philanthropic strategies and tactics for change: A concise framework*. Washington, DC: The Aspen Institute.
- Gamson, W.A. (2000). Framing social policy. *Nonprofit Quarterly*, 7(2). Retrieved January 30, 2001, from <http://www.newenglandnonprofit.org/advocacy/framing.htm>.
- Gamson, W.A., & Lasch, K.E. (1983). The political culture of social welfare policy. In S.E. Spiro and E. Yuchtman-Yaar (Eds.), *Evaluating the welfare state: Social and political perspectives* (pp. 397- 415). New York: Academic Press.
- Gerber, L., & Dorfman, L. (1996, August). Henry Horner Mothers Guild: Tenants go public on public housing. Bethesda, MD: University Research Corporation.
- Harvard Family Research Project (2001, June). *Media research findings in brief*. Cambridge, MA: Author.
- Harvard Family Research Project (2002, February). *Devolution Initiative learning report: 1997-2001 evaluation findings*. Cambridge, MA: Author.
- Heintz-Knowles, K. (1998). *Children and welfare reform: High stakes, low coverage*. Oakland, CA: Children Now.
- Henry, G.T. (2002). Ask the expert. *The Evaluation Exchange*, 8(3), 16.
- Henry, G.T., & Gordon, C.S. (2001). Tracking issue attention: Specifying the dynamics of the public agenda. *Public Opinion Quarterly*, 65, 157-177.
- Henry, G.T., & Gordon, C.S. (2003). Driving less for better air: Behavioral impacts of a public information campaign. *Journal of Policy Analysis and Management*, 22(1), 45-64.
- Henry, G.T., & Rivera, M. (1998, October). *Public information campaigns and changing behaviors*. Paper presented at the meeting of the Association for Public Policy Analysis and Management, New York.
- Holder, H.D., Gruenewald, P.J., Ponicki, W.R., Treno, A.J., Grube, J.W., Saltz, R.F., Voas, R.B., Reynolds, R., Davis, J. Sanchez, L., Guamont, G., & Roeper, P. (2000). Effect of community-based interventions on high-risk drinking and alcohol-related injuries. *Journal of the American Medical Association*, 284(18), 2341-2347.
- Holder, H.D., & Treno, A.J. (1997). Media advocacy in community prevention: News as a means to advance policy change. *Addiction*, 92, S189-S199.
- Iyengar, S. (1991). *Is anyone responsible? How television frames political issues*. Chicago: University of Chicago Press.

- Iyengar, S., & Kinder, D.R. (1987). *News that matters*. Chicago: University of Chicago Press.
- Kingdon, J. (1984). *Agendas, Alternatives, and Public Policies*. Boston, MA: Little Brown.
- Klein, E. (2002). Question and Answer. *The Evaluation Exchange*, 8(3), 10-11.
- McCombs, M., & Shaw, D.L. (1973). The agenda-setting function of the mass media. *Public Opinion Quarterly*, 37, 62-75.
- Ryan, C. (1991). *Prime time activism*. Boston, MA: South End Press.
- Ryan, C. (1996). The media war over welfare. *Peace Review*, 3(1), 13-19.
- Salmon, C.T, Post, L.A., & Christensen, R.E. (2003, March). *Mobilizing public will for social change*. Lansing, MI: Michigan State University.
- Sutton, S., & Thompson, E. (2002). Ask the expert. *The Evaluation Exchange*, 8(3), 17.
- Tabachnick, J. & Dawson, E. (2000). *STOP IT NOW! Report #5 - Four year evaluation: Findings reveal success of STOP IT NOW! VERMONT*. Haydenville, MA: STOP IT NOW.
- Treno, A.J., Breed, L., Holder, H., Roeper, P., Thomas, B.A., & Gruenewald, P.J. (1996). Evaluation of media advocacy efforts within a community trial to reduce alcohol-involved injury. *Evaluation Review*, 20(4), 404-423.
- Treno, A.J., & Holder, H.D. (1997a). Community mobilization: Evaluation of an environmental approach to local action. *Addiction*, 92, S173-S187.
- Treno, A.J., & Holder, H.D. (1997b). Community mobilization, organizing, and media advocacy: A discussion of methodological issues. *Evaluation Review*, 21(2), 166-190.
- Villarosa, L. (2002, December 3). To prevent sexual abuse, abusers step forward. *The New York Times*.
- Wallack, L., & Dorfman, L. (2001). Putting policy into health communication: The role of media advocacy. In R.E. Rice & C.K. Atkin (Eds.), *Public communication campaigns* (pp. 389-401). Thousand Oaks, CA: Sage.
- Weiss, J.A., & Tschirhart, M. (1994). Public information campaigns as policy instruments. *Journal of Policy Analysis and Management*, 13(1), 82-119.
- Weitzman, B.C., Silver, D. & Dillman, K. (2002). Integrating a comparison group design into a theory of change evaluation: The case of the Urban Health Initiative. *American Journal of Evaluation*, 23(4), 371-385.

APPENDIX A

Project to Build Evaluation Capacity of Nonprofit Communications First-Stage Research Projects

(Coordinated by the Communications Consortium Media Center)

Coffman, J. (2002, April). *Public communication campaign evaluation: An environmental scan of challenges, criticisms, practice, and opportunities*. Cambridge, MA: Harvard Family Research Project. Presents what has been happening in the field of public communication campaign evaluation in recent years. It examines evaluation challenges, criticisms, and practice and includes sections on relevant theory, outcomes, and useful methods for designing evaluations. It ends with opportunities for the road ahead.

Dorfman, L., Ervace J., Woodruff, K. (2002, November). *Voices for change: A taxonomy of public communications campaigns and their evaluation challenges*. Berkeley, CA: Berkeley Media Studies Group. Makes the case that communications campaigns cover a broad range of different types and characteristics, and can be differentiated along the axes of **purpose**, **scope**, and **maturity**. Examines what communications campaigns that fall on different areas of these three axes look like, and how *where* they fall may affect the evaluation approach used and lead to distinct evaluation challenges and needs.

Salmon, C.T, Post, L.A., & Christensen, R.E. (2003, June). *Mobilizing public will for social change*. Lansing, MI: Michigan State University. Examines the theory and strategies of “public will” campaigns and offers tangible criteria for their evaluation. It provides a rich inventory of strategies for use in mobilizing the public will through an integration of models of agenda building, social problem construction, issues management, social movements, media advocacy, and social capital. In addition, the paper provides cases and examples of public will campaigns directed at various social problems, along with criteria for evaluating these campaigns at various stages of a social problem’s life cycle.

Coffman, J. (2003, June). *Lessons in evaluating communications campaigns: Five case studies*. Cambridge, MA: Harvard Family Research Project. Builds on the findings of the first and second papers. It examines specifically how campaigns with different purposes (individual behavior change and policy change) have been evaluated, and how evaluators have tackled some of the associated evaluation challenges that the first three papers raised as important to address. It features five brief case studies in which the main unit of analysis is not the campaign, but the campaign’s *evaluation*. The case studies provide a brief snapshot of the real experiences of campaign evaluations. The paper also features cross-case lessons that highlight important findings and themes.